Attorney's Docket No.: 881-011446-US (PAR) **PATENT**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: ABUTALEBI et al
Serial No.: 10/642,847
Filed: August 18, 2003
For: Method And System For Processing Subband Signals Using Adaptive Filters

Group No.: 2643

Examiner:

RECEIVED
FEB 0 3 2004
Technology Center 2600

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF CERTIFIED COPY

Attached please find the certified copy of the foreign application from which priority is claimed for this case:

Country	:	**Canada**
Application Number	:	**2,399,159**
Filing Date	:	**August 16, 2002**

FEB 0 4 2004

WARNING: "When a document that is required by statute to be certified must be filed, a copy, including a photocopy or facsimile transmission of the certification is not acceptable." 37 CFR 1.4(f) (emphasis added.)

Geza Ziegler

SIGNATURE OF ATTORNEY

Reg. No.: 44,004

Geza C. Ziegler, Jr.

Type or print name of attorney

Tel. No.: (203) 259-1800

Perman & Green, LLP

Customer No.: 2512

P.O. Address

425 Post Road, Fairfield, CT 06824

NOTE: The claim to priority need be in no special form and may be made by the attorney or agent if the foreign application is referred to in the oath or declaration as required by § 1.63.

CERTIFICATE OF MAILING/TRANSMISSION (37 CFR 1.8a)

I hereby certify that this correspondence is, on the date shown below, being:

MAILING

☒ deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

FACSIMILE

☐ transmitted by facsimile to the Patent and Trademark Office

Alicia Czako

Signature

Date: January 28, 2004

Alicia J. Czako
(type or print name of person certifying)

(Transmittal of Certified Copy [5-4]

Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

Bureau canadien des brevets

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

Canadian Patent Office

Certification

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No:**2,399,159**, on August 16, 2002, by **DSPFACTORY LTD.**, assignee of Hamid Reza Abutalebi, Robert Brennan, Hamid Sheikhzadeh Nadjar and Dequn Sun, for "Convergence Improvement for Oversampled Subband Adaptive Filters".

L. Régimbal

Agent certificateur/Certifying Officer

November 19, 2003

Date

Canada

(CIPO 68)
04-09-02

OPIC CIPO

Abstract of the Disclosure

A method of improving the convergence properties of the oversampled subband adaptive filters is disclosed. The method comprises steps of: (a) whitening by spectral emphasis, where, after WOLA analysis, subband signals are decimated by a factor of M/OS where M is the number of filters and OS is the oversampling factor; or (b) whitening by additive noise, where high-pass noise is added to bandpass signals to make them whiter in spectrum; or (c) whitening by decimation, where the subband signals are further decimated by a factor of DEC<OS; or (d) a combination of the above steps (a), (b) and (c).

Convergence Improvement for Oversampled Subband Adaptive Filters

Field of the Invention

The present invention relates to convergence Improvement techniques for oversampled subband adaptive filters.

Background of the Invention

It is well known that a noise cancellation system can be implemented with a fullband adaptive filter working on the entire frequency band of interest [4]. The Least Mean-Square (LMS) algorithm and its variants are often used to adapt the fullband filter with relatively low computation complexity and good performance. However, the fullband LMS solution suffers from significantly degraded performance with colored interfering signals due to large eigenvalue spread and slow convergence [4,5,6]. Moreover, as the length of the LMS filter is increased, the convergence rate of the LMS algorithm decreases and computational requirements increase. This can be a problem in applications, such as acoustic echo cancellation, that demand long adaptive filters to model the return path response and delay. These issues are especially important in portable applications, where processing power must be conserved.

As a result, subband adaptive filters (SAF) become a viable option for many adaptive systems. The SAF approach uses a filterbank to split the fullband signal input into a number of frequency bands, each serving as input to an adaptive filter. The subband decomposition greatly reduces the update rate and the length of the adaptive filters resulting in a much lower computational complexity. Further, subband signals are often decimated in SAF systems. This leads to a whitening of the input signals and an improved convergence behavior [7]. If critical sampling is employed, the presence of aliasing distortions requires the use of adaptive cross-filters between adjacent subbands or gap filterbanks [7,8]. However, systems with cross-filters generally converge slower and have higher computational cost, while gap filterbanks produce significant signal distortion. Oversampled SAF systems offer a simplified structure that without employing cross-filters or gap filterbanks, reduce the alias level in subbands. To

reduce the computation cost, often a close to one non-integer decimation ratio is used [9].

Summary of the Invention

The inventors have investigated the convergence properties of an SAF system based on generalized DFT (GDFT) filterbanks. The filterbank is a highly oversampled one (oversampling by a factor of 2 or 4 or more). Due to the ease of implementation, low-group delay and other application constraints we chose a higher oversampling ratio than those typically proposed in the literature.

The oversampled input signals received by the subband processing blocks are no longer white in spectrum. In fact, for oversampling factors of 2 and 4, their bandwidth will be limited to $\pi/2$ and $\pi/4$ respectively. As a result, one would expect a slow convergence rate due to eigenvalue spread problem [4,5,6]. On the other hand, while the oversampled subband signals are not white, their spectra are colored in a predicable way and can therefore be modified by further processing to "whiten" them in order to increase the convergence rate. Thus, the inherent benefit of decreased spectral dynamics resulting from subband decomposition is not lost due to oversampling. Various spectral whitening techniques will be described hereafter. Another method of improving the convergence rate is to employ adaptation strategies that are less sensitive to eigenvalue spread problem. One of these strategies is the Affine Projection (AP) algorithm. Exact and approximate versions of the AP algorithm are proposed to speed up the convergence rate of the SAF system on an oversampled filterbank.

A further understanding of other features, aspects and advantages of the present invention will be realized by reference to the following description, appended claims, and accompanying drawings.

Brief Description of the Drawings

A preferred embodiment(s) of the invention will be described with reference to the accompanying drawings, in which:

Figure 1 shows a block diagram of whitening by spectral emphasis method;

Figure 2 shows a block diagram of whitening by additive noise method;

Figure 3 shows a block diagram of whitening by decimation method;

Figure 4 shows a signal spectra at various points of Figure 3; and

Figure 5 shows Average Normalized Filter MSE for speech in 0 dB SNR White noise, (a) without whitening, (b) whitening by spectral emphasis, (c) whitening by decimation..

Figure 6 shows eigenvalues of the autocorrelation matrix of the reference signal for: No whitening, Whitening by spectral emphasis, whitening by decimation, and whitening by decimation and spectral emphasis.

Figure 7 shows measured mean-squared error for: No whitening, whitening by spectral emphasis, Whitening by decimation, and whitening by decimation and spectral emphasis.

Figure 8 shows measured mean-squared error for APA with different orders

Detailed Description of the Preferred Embodiment(s)

Whitening by spectral emphasis

Figure 1 shows a block diagram of an SAF system that includes the proposed "whitening by spectral emphasis" method. As shown an unknown plant P(z) is modeled by the adaptive filter, W(z). After WOLA analysis, subband signals are decimated by a factor of M/OS, where M is the number of filters, and OS is the oversampling factor. At this stage, the subband signals are no longer full-band. Rather, as shown in Figure 1 (points 1 and 2), their bandwidth is now π/OS. The emphasis filter ($g_{pre}(z)$) then amplifies the high frequency contents of signals at points 1 and 2 to obtain almost white spectra. The filter gain (G) is a design parameter that depends on the analysis filter shape.

Whit ning by additive n ise

Alternatively, high-pass noise can be added to bandpass signals to make them whiter in spectrum. As shown in Figure 2, first the average power (G) of the signal at point 1 is estimated and used to modulate a high-pass noise a(n). The input to adaptive filter (point 3) is then whitened by adding G.a(n) to the signal at point 1.

Whitening by decimation

Figure 3 shows a block diagram of the SAF system with a proposed "whitening by decimation" method. As shown, the subband signals (for both the reference input x(n) and the primary input d(n)) are further decimated by a factor of DEC<OS. Assume, without loss of generality, that DEC is at its maximum, DEC= OS-1. As demonstrated in Figure 4 (point 3), this increases the bandwidth to π (OS-1)/OS ($3\pi/4$ for OS=4) without generating in-band aliasing. Due to the increased bandwidth, the LMS algorithm now converges much faster. To be able to use the adaptive filter ($W_d(z)$), it should be expanded by OS-1. This creates in-band images (point 4 in Figure 4). However, since the signal at point 1 does not contain considerable energy for $\omega > \pi/OS$, the spectral images will not contribute to any errors.

Affine Projection

In order to further increase the convergence rate, a class of adaptive algorithms called Affine Projection have been proposed [12]. Affine Projection Algorithm (APA) forms a link between Normalized LMS (NLMS) and Recursive Least Square (RLS) adaptation algorithms: faster convergence of RLS and low computational requirements of NLMS are compromised in APA.

In NLMS, the new adaptive filter weights have to best fit the last input vector to the corresponding desired signal. In APA, this fitting expands to the **P-1** past input v ctors (**P** being the APA order). Adaptation algorithm for the **P**th order APA can be summarized as follows:

1) update X_n and d_n

2) $e_n = d_n - X_n^T W_n^*$

3) $W_{n+1} = W_n + \mu X_n (X_n^H X_n + \alpha I)^{-1} e_n^*$

where:

X_n: an $L \times P$ matrix containing **P** past input vectors

d_n: a vector of the past **P** past desired signal samples

W_n: adaptive filter weights vector at time n

α : regularization factor

The convergence of APA is surveyed in [12, 13]. It is shown that as projection order **P** increases, the convergence rate becomes less dependant on the eigenvalue spread. Increasing the APA order results in faster convergence at the cost of more computational complexity of the adaptation algorithm.

We propose the use of the APA for a SAF system implemented on a highly oversampled WOLA filterbank [1,2,3]. An APA order of **P** = 2 can be a good choice, compromising fast convergence and low complexity. In this case, the matrix $X_n^H X_n$ can be approximated by R (autocorrelation matrix of the reference signal) [14]. So, for **P** = 2, it is sufficient to estimate the first two autocorrelation coefficients (r(0) and r(1)) and then inverse the matrix R, analytically. A first order recursive smoothing filter can be used to estimate r(0) and r(1).

Combination of the above techniques

It is possible to combine any two or more of the described techniques to achieve a high r performance. For example, whitening by decimation improves

the convergence rate by increasing the effective bandwidth of the reference signal. However, it cannot deal with the smallest eigenvalues that are associated with the stop band region of the analysis filter. On the other hand, whitening by spectral emphasis improves the convergence by limiting the stop band loss thereby increasing the smallest eigenvalues. A combination of the two techniques will enable us to take advantage of the merits of both systems.

Performance evaluation

Preliminary assessments show that the performance of the whitening by additive noise is very similar to whitening by spectral emphasis. However, the computation cost of whitening by additive noise is less since it does not need emphasis filters. Instead, it needs a very simple filter (per subband) to estimate the signal power.

Figure 5 shows typical convergence behavior of the proposed whitening by decimation compared to no whitening and whitening by emphasis. The application of the SAF system has been 2-microphone adaptive noise cancellation. As shown, whitening by decimation converges mush faster than the other two methods.

Whitening by decimation greatly improves the convergence properties of the SAF system. At the same time, since the adaptive filter operates at a low frequency, the method offers less computation than whitening by emphasis or by adding noise. However, the proposed whitening by decimation is applicable only to oversampling factors (OS) of more than 2. For detailed mathematical models of SAF systems see [9,15].

Figure 6 shows the theoretical Eigenvalues of the autocorrelation matrix of the reference signal for: No whitening, Whitening by spectral emphasis, Whitening by decimation, and Whitening by decimation and spectral emphasis. The method employed is described in [6]. As shown, while whitening by spectral emphasis and by decimation both offer improvements (demonstrated by a rise in th eigenvalues), a combination of both method is more promising. This conclusion is confirmed by the mean-squared error (MSE) results shown in Figure 7. Finally, Figure 8 shows the MSE results APA orders of **P** = 1, 2, 4 and

5 (The APA for **P** = 1 yields an NLMS system). As shown, increasing the AP order, improves both the convergence rate and the MSE.

The present invention will be further understood by the additional description A, B and C attached hereto.

While the present invention has been described with reference to specific embodiments, the description is illustrative of the invention and is not to be construed as limiting the invention. Various modifications may occur to those skilled in the art without departing from the true spirit and scope of the invention as defined by the appended claims.

References

[1] R. Brennan and T. Schneider, "Filterbank Structure and Method for Filtering and Separating an Information Signal into Different Bands, Particularly for Audio Signal in Hearing Aids". *United States Patent 6,236,731. WO 98/47313*. April 16, 1997.

[2] R. Brennan and T. Schneider, "A Flexible Filterbank Structure for Extensive Signal Manipulations in Digital Hearing Aids", *Proc. IEEE Int. Symp. Circuits and Systems*, pp.569-572, 1998.

[3] R. Brennan and T. Schneider, "Apparatus for and method of filtering in an digital hearing aid, including an application specific integrated circuit and a programmable digital signal processor", *United States Patent 6,240,192, May 2001.*

[4] B. Widrow et al., "Adaptive noise cancellation: Principles and applications". *Proc. IEEE*, vol. 63, no. 12, Dec. 1975.

[5] Haykin, S., *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 3rd Edition, 1996.

[6] Dennis R. Morgan, "Slow Asymptotic Convergence of LMS Acoustic Echo Cancelers", *IEEE Trans. Speech and Audio Proc.*, Vol. 3, No. 2, pp. 126-136, March 1995.

[7] A. Gilloire and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments and Applications to Acoustic Echo Cancellation". *IEEE Trans. Signal Processing*, vol. SP-40, no. 8, pp. 1862-1875, Aug. 1992.

[8] J. J. Shynk, "Frequency-Domain and Multirate Adaptive Filtering". *IEEE Signal Processing Magazine*, pp. 14-37, Jan. 1992.

[9] S. Weiss, "On Adaptive Filtering in Oversampled Sub-bands", PhD. Thesis, Signal Processing Division, University of Strathclyde, Glasgow, May 1998.

[10] King Tam et. al., "Sub-band Adaptive Signal Processing in an Oversampled Filterbank", IDF filed on August 7, 2002, Application No. 2,354,808.

[11] King Tam, Hamid Sheikhzadeh, and Todd Schneider, "highly oversampled subband adaptive filters for noise cancellation on a low-resource dsp system", *Proc. Of ICSLP 2002.*

[12] K. Ozeki and T. Umeda, "An adaptive algorithm filtering using an orthogonal projection to the affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, no. 5, pp.19-27, Feb. 1984.

[13] M. Montazeri and P. Duhamel, "A set of algorithms linking NLMS and block RLS algorithms" *IEEE Tran. on Signal Processing*, vol. 43, no. 2, pp. 444-453, Feb. 1995.

[14] V. Myllyla, "Robust fast affine projection algorithm for acoustic echo cancellation," *in proc. of Inter. Workshop on Acoustic Echo and Noise Control*, Sep. 2001.

[15] S. Weiss et al., "Polyphase Analysis of Subband Adaptive Filters", 33rd Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, 1999.

Additional D scription A

Technical R port

"Polyphase Analysis of Subband Adaptive Filters"

9-1

Polyphas Analysis of Subband Adaptive Filters

Stephan Weiss[1], Robert W. Stewart[2], Moritz Harteneck[3], and Alexander Stenger[4]

[1] Dept. Electronics & Computer Science, University of Southampton, UK
[2] Dept. Electronic & Electrical Eng., University of Strathclyde, Glasgow, UK
[3] Infineon Technologies AG, Munich, Germany
[4]Telecommunications Institute I, University of Erlangen, Germany
s.weiss@ecs.soton.ac.uk, r.stewart@eee.strath.ac.uk

Abstract

Based on a polyphase analysis of a subband adaptive filter (SAF) system, it is possible to calculate the optimum subband impulse responses to which the SAF system will converge. In this paper, we give some insight into how these optimum impulse responses are calculated, and discuss two applications of our technique. Firstly, the performance limitations of an SAF system can be explored with respect to the minimum mean square error performance. Secondly, fullband impulse responses can be correctly projected into the subband domain, which is required for example for translating constraints for subband adaptive beamforming. Examples for both applications are presented.

1. Introduction

Adaptive filtering in subbands is a popular approach to a number of problems, where high computational cost and slow convergence due to long filters permits the direct implementation of a fullband algorithm. These problems include acoustic echo cancellation [5, 3], identification of room acoustics [8], equalization of acoustics [10], or beamforming [6, 11]. In Fig. 1, a subband adaptive filter (SAF) is shown in a system identification setup of an unknown system $s[n]$, whereby the input $x[n]$ and the desired signal $d[n]$ are split into K frequency bands by analysis filter banks built of bandpass filters $h_k[n]$. Assuming a cross-band free SAF design [3], an adaptive filter $w_k[n]$ is applied to each subband decimated by $N \leq K$. Finally, the fullband error signal e[n] can be reconstructed via a synthesis bank.

However, subband adaptive filters (SAF) are subject to a number of limitations, which have been investigated, for example, with respect to the required filter length [3, 14] or to lower bounds for the MMSE and the modelling accuracy [12]. These analyses have been performed using modulation description [3, 7], time domain [14], or frequency domain approaches [5, 12].

Here, we discuss the SAF in Fig. 1 using a polyphase description of the signals and filters therein [2]. This will provide some new and alternative insight into the optimality of SAFs. Sec. 2 analyses the subband errors, which leads to the derivation and discussion of an optimal subband filter structure in Sec. 3. Application examples for the proposed techniques are underlined by simulations in Sec. 4.

2. Polyphase Analysis of Subband Errors

The aim of this section is to express the subband error signals, $E_k^d(z) \bullet\!\!-\!\!\circ\, e_k(z)$, in terms of the polyphase components of all involved signals and systems. Implicitly, this means that we are trying to find a linear, time-invariant (LTI) description of the error signal. To achieve this task, we first require suitable representations for the decimated desired signal in the kth subband, $D_k^d(z) \bullet\!\!-\!\!\circ\, d_k[n]$, and for the decimated input signal in the kth subband, $X_k^d(z) \bullet\!\!-\!\!\circ\, x_k[n]$, as labelled in Fig. 1. In our notation, superscript $\{\cdot\}^d$ for z-transforms of signals refers to decimated quantities, while normal variables such as $X_k(z)$ indicate undecimated signals, i.e. in this case the input signal in the kth subband before going into the decimator as shown in Fig. 1.

The formulation of the kth decimated desired signal $D_k^d(z) \bullet\!\!-\!\!\circ\, d_k[n]$ will be the first aim. We define the expansion of the desired signal $D(z) \bullet\!\!-\!\!\circ\, d[n]$ into

9-2

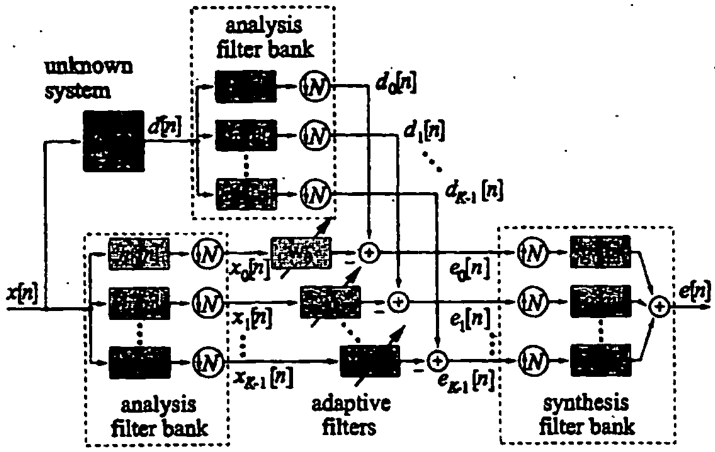


Fig. 1. Subband adaptive filter (SAF) in a system identification setup.

type-2-polyphase components [9] $D_n(z)$,

$$D(z) = \sum_{n=0}^{N-1} z^{-N+n+1} \cdot D_n(z^N) , \qquad (1)$$

and a type-1-polyphase expansion [9] of the analysis filters $H_k(z)$,

$$H_k(z) = \sum_{n=0}^{N-1} z^{-n} \cdot H_{k,n}(z^N) . \qquad (2)$$

Similarly, for all following polyphase expansions, it is assumed for compatibility that systems are represented by a type-1-polyphase expansion, and signals by type-2-polyphase expansions. Bringing these polyphase components of (1) and (2) into vector form,

$$\underline{D}(z) = [D_0(z)\ D_1(z)\ \cdots\ D_{N-1}(z)]^T \qquad (3)$$

$$\underline{H}_k(z) = [H_{k|0}(z)\ H_{k|1}\ \cdots\ H_{k|N-1}(z)]^T \qquad (4)$$

$D_k^d(z)$ can now be expressed as

$$D_k^d(z) = \underline{H}_k^T(z) \cdot \underline{D}(z) \quad . \qquad (5)$$

To trace the desired signal back to the input signal $X(z) \bullet\!-\!\circ\, x[n]$, the expression $D(z) = S(z) \cdot X(z)$ can be appropriately expanded such that the nth polyphase component in (3) may be written as

$$D_n(z) = \underline{S}^T(z) \cdot \Lambda_n(z) \cdot \underline{X}(z) \quad . \qquad (6)$$

The vector $\underline{S}(z)$ contains the type-1-polyphase components of the unknown system $S(z) \bullet\!-\!\circ\, s[n]$, while $\underline{X}(z)$ is defined similarly to (3) based on the type-2-polyphase components of the input signal $X(z) \bullet\!-\!\circ\, x[n]$. The matrix $\Lambda_n(z)$ in (6) is a delay matrix defined as

$$\Lambda_n(z) = \begin{bmatrix} 0 & \mathbf{I}_{N-n} \\ z^{-1}\mathbf{I}_n & 0 \end{bmatrix} . \qquad (7)$$

With (5) and (6), the decimated kth desired subband signal $D_k^d(z)$

$$D_k^d(z) = \underline{H}_k^T(z) \begin{bmatrix} \underline{S}^T(z)\Lambda_0(z) \\ \underline{S}^T(z)\Lambda_1(z) \\ \vdots \\ \underline{S}^T(z)\Lambda_{N-1}(z) \end{bmatrix} \underline{X}(z) = \underline{H}_k^T(z)\mathbf{S}(z)\underline{X}(z) \qquad (8)$$

can be assembled. For brevity, the substituted matrix $\mathbf{S}(z)$ holds differently delayed polyphase components of the unknown system.

With the type-2-polyphase components of $X(z)$ and the polyphase representation of the analysis filter bank in (2) it is comparably simple to derive the kth decimated input signal $X_k^d(z)$ as

$$X_k^d(z) = \underline{H}_k^T(z) \cdot \underline{X}(z) \quad . \qquad (9)$$

Finally, with (8), (9), and the transfer function of the kth adaptive filter $W_k(z) \bullet\!-\!\circ\, w_k[n]$ it is possible to formulate the kth subband error signal, $E_k^d(z) \bullet\!-\!\circ\, e_k[n]$:

$$E_k^d(z) = D_k^d(z) - W_k(z) \cdot X_k^d(z) \qquad (10)$$

$$= \left\{ \underline{H}_k^T(z) \cdot \mathbf{S}(z) - \underline{H}_k^T(z) \cdot W_k(z) \right\} \underline{X}(z) \qquad (11)$$

9-3

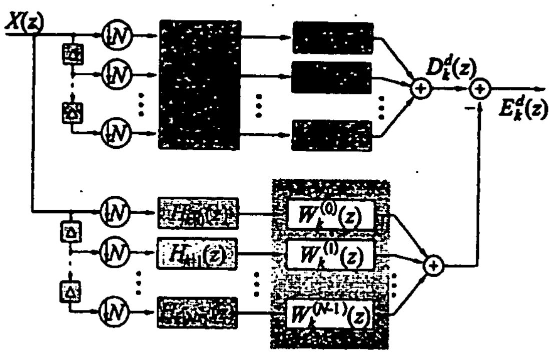


Fig. 2. SAF optimal polyphase solutions in the kth subband.

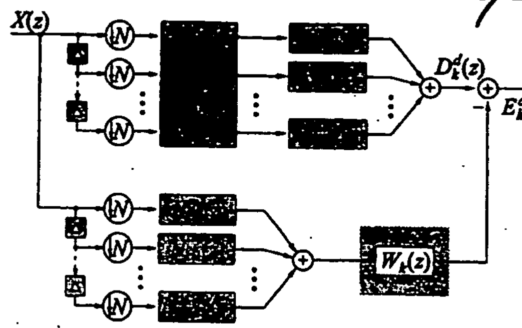


Fig. 3. SAF standard solution in the kth subband.

Note, that for the description of $E_k^d(z)$, the time-varying decimators have been swapped with all system elements in the SAF structure of Fig. 1, and (11) only contains LTI terms.

3. Subband Error Minimization

This section discusses the optimum subband filters to solve the identification problem outlined in Sec. 1, based on the polyphase analysis of the subband errors in the previous Sec. 2.

3.1. Optimum Subband Filters

As no external disturbance of the SAF system in Fig. 1 by observation noise is present, ideally the attainable minimum mean square error (MMSE) should be zero. This is identical to setting $E_k^d(z)$ in (11) equal to zero. As independence of the optimum solution from the input signal's polyphase components in $\underline{X}(z)$ is desirable, the requirement for optimality (in every sense) is given by

$$\underline{H}_k^T(z) \cdot \mathrm{S}(z) \stackrel{!}{=} \underline{H}_k^T \cdot W_{k,\mathrm{opt}}(z) \quad . \tag{12}$$

Hence, we obtain N cancellation conditions indicated by superscripts $\{\cdot\}^{(n)}$, which have to be fulfilled:

$$W_{k,\mathrm{opt}}^{(n)}(z) = \frac{\underline{H}_k^T(z) \cdot \mathbf{\Lambda}_n^T(z) \cdot \underline{S}(z)}{H_{k|n}(z)} \qquad \forall n \in \{0; N-1\} . \tag{13}$$

Therefore, ideally $W_k(z)$ in (11) and (12) should be replaced by an $N \times N$ diagonal matrix with entries $W_k^{(n)}(z)$, $n = 0(1)N-1$. For the kth subband, this solution with N polyphase filters is depicted in Fig. 2.

3.2. Interpretation

Alternatively, the nth optimum solution can be written as

$$W_{k,\mathrm{opt}}^{(n)}(z) = \sum_{\nu=0}^{N-1} A_{k|n}^{(\nu)}(z) \cdot S_\nu(z) \quad . \tag{14}$$

and interpreted as a superposition of polyphase components $S_\nu(z)$ of the unknown system $S(z)$, "weighted" by transfer functions

$$A_{k|n}^{(\nu)}(z) = z^{-\lfloor (n+\nu)/N \rfloor} \cdot \frac{H_{k|(n+\nu) \bmod N}(z)}{H_{k|n}(z)} . \tag{15}$$

From this, we can observe, that the length of the optimum subband responses is obviously given by $1/N$ of the order of $S(z)$, but extended by the transfer functions (15). These extending transients are causal for poles of $A_{k|n}^{(\nu)}(z)$ within the unit circle, and acausal for stabilized poles outside the unit-circle [13], motivating a non-causal optimum response.

Further, for an ideal, alias-free filter bank, the polyphase components $H_{k|n}(z)$ in (15) should not differ in magnitude but only in phase, which is compensated for by the delay element in (15). Hence all N solutions become identical, an the N optimum polyphase filters can be replaced by a single filter $W_{k,\mathrm{opt}}(z)$ as shown in Fig. 3, which is equivalent to the original standard setup in Fig. 1. In general, and particularly if aliasing is present, the optimum polyphase solutions $W_{k,\mathrm{opt}}^{(n)}(z)$ will differ. In this case the optimum standard SAF solution according to Fig. 3 gives the closest l_2 match to all N polyphase solutions:

$$W_{k,\mathrm{opt}}(z) = \frac{1}{N} \sum_{n=0}^{N-1} W_{k,\mathrm{opt}}^{(n)}(z) \quad . \tag{16}$$

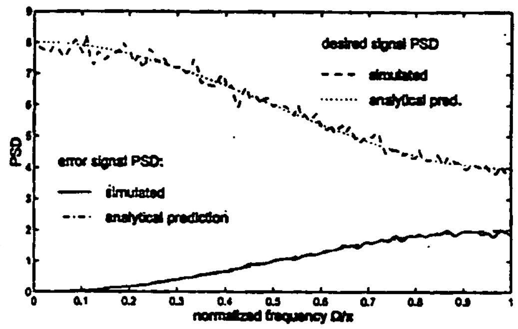


Fig. 4. Comparison between simulated and analytically predicted PSDs in the 0th subband.

The error made in this approximation explains error and modelling limitations of the SAF approach and represents an alternative coefficient / time-domain description as opposed to spectrally motivated SAF error explanations in the literature [3, 12]. Interestingly, in [7] the same polyphase structure as in Fig. 2 is obtained using modulation description [2] 5 , although only for the critically sampled case.

4. Applications and Simulations

We now want to explore two applications for the polyphase analysis presented in Secs. 2 and 3.

4.1. Error Limits

A very basic example given in the following will demonstrate the ability of the proposed analysis to predict optimal subband responses and error terms in the context of SAF systems. For this example, a 2-channel critically decimated standard SAF system as in Fig. 1 based on a Haar filter bank [2] should adaptive identify an unknown system $S(z) = 1 + z^1$, using unit variance Gaussian white noise excitation. Looking at the channel $k = 0$ produced by the Haar lowpass filter $H_0(z) = 1 + z^{-1}$, an RLS adaptive algorithm [4] converges to the solution

$$W_{0,\mathrm{adapt}}(z) = 1.4873 + 0.5067z^{-1} \quad . \qquad (17)$$

Analytical evaluation via (14) and (15) yields the $N = 2$ optimum polyphase solutions for the band $k = 0$

$$W^{(0)}_{0,\mathrm{opt}}(z) = 2 \ ; \quad W^{(1)}_{0,\mathrm{opt}}(z) = 1 + z^{-1} \ , \qquad (18)$$

which refers to the optimal subband adaptive filter structure shown in Fig. 2. If this setup is simplified to the structure of the standard SAF system in Fig. 3, the analytical solution (16) calculated from (18) is given by the mean of the two optimum polyphase solutions,

$$W_{0,\mathrm{opt}}(z) = 1.5 + 0.5z^{-1} \quad .$$

This result obviously very closely agrees with the simulation result in (17).

Based on the above analytical solutions, it is now possible to predict the subband error signal as due to the mismatch of (18) and (4.1). The PSD of the 0th adapted subband error signal, $S_{e_0}(e^{j\Omega})$, can be analytically predicted by inserting the optimum standard solution (16) into (11),

$$S_{e_0}(e^{j\Omega}) = |E^d_k(e^{j\Omega})|^2 = 1 - \cos\Omega \quad , \qquad (19)$$

which can be used to determine the minimum mean squared error of the SAF system alternative to spectral methods [12]. Fig. 4 demonstrates the excellent fit between the analytically calculated PSD in (19), and the measured results from the RLS simulation. Also shown is the analytically predicted and measured PSD of the 0th desired subband signal $S_{d_0}(e^{j\Omega}) = 6 + 2\cos\Omega$ (hence the uncancelled error signal) calculated via (5).

4.2. Subband Projection

A second application example is concerned with substituting subband adaptive system identification with the proposed analysis. If a digital impulse response is given in the fullband, but should be projected into the subband domain, an SAF identification is mostly required. This could be to produce computationally efficient sound processing from a given (fullband) room transfer function [8], or the projection of constraints into the subband domain when performing subband adaptive beamforming [11].

We assume an SAF system with $K = 8$ channels decimated by $N = 6$, and wide analysis filters to improve spectral whitening in the subbands [1]. Analysis and synthesis banks are derived from the two different prototype filters shown in Fig. 5. With a lowpass fullband response $s[n]$ given, an RLS adaptive identification yields in the subband $k = 0$ the coefficients shown in Fig. 6, along with the analytic solution according to (14) and (16). For the analytic solution, the roots of the denominator polynomial in (15) have been substituted by appropriate causal and a-causal FIR filters. Obviously, the match between adaptive and analytical solution is very close, and themore direct analytical approach can replace an adaptive projection.

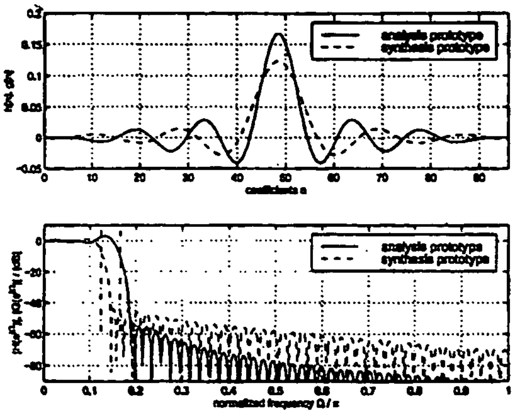


Fig. 5. Prototype filters.

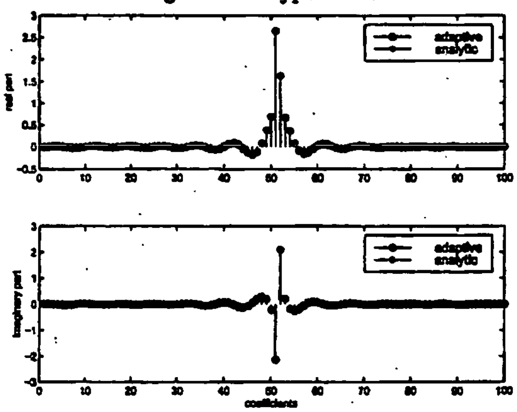


Fig. 6. Adaptive and analytic subband response for $k = 0$.

5. Conclusions

We have introduced an analysis of an SAF system, which formulates the subband errors in dependency of LTI polyphase components only. The main result was a structural difference between what the optimum SAF requires and what the standard SAF structure provides. As a qualitative measure, this difference in structure gives alternative insight into the inaccuracies and limitations of the standard SAF approach. But as demonstrated, the approach can also be utilized to quantify errors. Different from alias measurement methods for error prediction [12], the analysis also offers access to the coefficient domain and thus allows us to state optimum SAF subband responses. As an application for the latter, an example was given that allows us to substitute the subband projection by SAF system identification with the proposed analytical polyphase approach.

6. Acknowledgements

The authors gratefully acknowledge Dr. Ian K. Proudler, of DERA, Malvern, UK, who partially supported this work. S. Weiss would like to thank the Royal Academy of Engineering for providing a travel grant.

References

[1] P.L. de León II and D.M. Etter. "Experimental Results with Increased Bandwidth Analysis Filters in Oversampled Subband Acoustic Echo Cancelers". *IEEE Sig Proc Letters*, Vol.2(No.1):pp.1–3, Jan. 1995.

[2] N.J. Fliege. *Multirate Digital Signal Processing: Multirate Systems, Filter Banks, Wavelets.* Wiley, 1994.

[3] A. Gilloire and M. Vetterli. "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments and Applications to Acoustic Echo Cancelation". *IEEE Trans Signal Processing*, Vol.40 (No.8):pp.1862–1875, Aug. 1992.

[4] S. Haykin. *Adaptive Filter Theory.* Prentice Hall, 2nd ed, 1991.

[5] W. Kellermann. "Analysis and Design of Multirate Systems for Cancellation of Acoustical Echoes". In *Proc. ICASSP*, vol 5, pp.2570–2573, New York, 1988.

[6] W. Kellermann. "Strategies for Combining Acoustic Echo Cancellation and Adaptive Beamforming Microphone Arrays". In *Proc. IEEE ICASSP*, vol I, pp.219–222, Munich, April 1997.

[7] S. S. Pradhan and V. U. Reddy. "A New Approach to Subband Adaptive Filtering". *IEEE Trans Signal Processing*, Vol.47(No.3):pp.655–664, March 1999.

[8] M. Schönle, N.J. Fliege, and U. Zölzer. "Parametric Approximation of Room Impulse Responses by Multirate Systems". In *Proc. IEEE ICASSP*, vol I, pp.153–156, Minneapolis, May 1993.

[9] P.P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice Hall, 1993.

[10] S. Weiß, S.R. Dooley, R.W. Stewart, and A.K. Nandi. "Adaptive Equalization in Oversampled Subbands". *IEE Elec. Let.*, Vol.34(No.15):pp.1452–1453, July 1998.

[11] S. Weiss, R.W. Stewart, M. Schabert, I.K. Proudler, and M.W. Hoffman. "An Efficient Scheme for Broadband Adaptive Beamforming". In *Asilomar Conf Sig. Sys. Comp.*, Monterey, CA, Nov. 1999.

[12] S. Weiß, R.W. Stewart, A. Stenger, and R. Rabenstein. "Performance Limitations of Subband Adaptive Filters". In *Proc. EUSIPCO*, vol. III, pp. 1245–1248, Rodos, Sept. 1998.

[13] B. Widrow and E. Walach. *Adaptive Inverse Control.* Prentice Hall, 1995.

[14] R.J. Wilson, P.A. Naylor, and D. Brookes. "Performance Limitations of Subband Acoustic Echo Controllers". In *Proc. IWAENC*, pp.176–179, London, Sept. 1997.

Additional Description B

Technical Report

"Highly Oversampled Subband Adaptive Filters for Noise Cancellation on a Low-Resource DSP System"

10-1

HIGHLY OVERSAMPLED SUBBAND ADAPTIVE FILTERS FOR NOISE CANCELLATION ON A LOW-RESOURCE DSP SYSTEM

King Tam, Hamid Sheikhzadeh, Todd Schneider

Dspfactory Ltd., 80 King Street South, Suite 206, Waterloo, Ontario, Canada N2J 1P5
e-mail: {King.Tam, hsheikh, Todd.Schneider}@dspfactory.com

ABSTRACT

A real-time subband adaptive noise cancellation system on an ultra low-power miniature DSP system is implemented. The system is targeted at personal communication devices where the speaker may be in a noisy environment. The system is implemented on an ultra low-power DSP system that incorporates a DSP core and an oversampled WOLA filterbank. Pre-emphasis filters are used to increase the convergence rate of a leaky LMS algorithm in the oversampled subband implementation. System performance is also improved relative to a fullband implementation due to benefits arising from using subband adaptive filters instead of fullband filters. A 10 dB reduction of noise power is achieved in tests using various noise conditions. The entire DSP system consumes 2.1 mW and can be realized in a package size of 6.5 x 3.5 x 2.5 mm.

1. INTRODUCTION

The objective of this research is to implement a subband adaptive noise cancellation system on an ultra low-power, small size, and low-cost platform. The system is targeted for telecommunication (e.g., headsets or mobile phones) or mobile speech recognition applications, where the user is talking in the presence of interfering noise. A robust system should provide significant noise cancellation, fast algorithmic convergence in colored noises, short group delay, and minimal introduction of artifacts into the speech signal. Furthermore, it should have low computational cost and complexity, low memory usage, low power requirements, and small physical size.

It is well known that a noise cancellation system can be implemented with a fullband adaptive filter working on the entire frequency band of interest [1]. The Least Mean-Square (LMS) algorithm and its variants are often used to adapt the fullband filter with relatively low computation complexity and good performance. However, the fullband LMS solution suffers from significantly degraded performance with colored interfering signals due to large eigenvalue spread and slow convergence [2]. Moreover, as the length of the LMS filter is increased, the convergence rate of the LMS algorithm decreases and computational requirements increase. This can be a problem in applications, such as acoustic echo cancellation, that demand long adaptive filters to model the return path response and delay. These issues are especially important in portable applications, where processing power must be conserved.

As a result, subband adaptive filters (SAF) become a viable option for many adaptive systems. The SAF approach uses a filterbank to split the fullband signal input into a number of frequency bands, each serving as input to an adaptive filter. The subband decomposition greatly reduces the update rate and the length of the adaptive filters resulting in a much lower computational complexity. Further, subband signal are often decimated in SAF systems. This leads to a whitening of the input signals and an improved convergence behavior [3]. If critical sampling is employed, the presence of aliasing distortions requires the use of adaptive cross-filters between adjacent subbands or gap filterbanks [3,4]. However, systems with cross-filters generally converge slower and have higher computational cost, while gap filterbanks produce significant signal distortion. Oversampled SAF systems offer a simplified structure that without employing cross-filters or gap filterbanks, reduce the alias level in subbands. To reduce the computation cost, often a close to one non-integer decimation ratio is used [5].

In this research we propose a SAF system based on generalized DFT (GDFT) filterbanks. The filterbank is a highly oversampled one (oversampling by a factor of 2 or 4). Due to the ease of implementation, low-group delay and other application constraints (explained in Section 3), we chose a higher oversampling ratio than those typically proposed in the literature. The convergence behavior due to the high oversampling rate is analyzed and properly addressed. An LMS-based version of the proposed SAF system is implemented on a DSP system that includes an oversampled filterbank. The DSP system [6,7] has a configurable oversampling rate of 2 or 4. The added computational cost due to sampling the subband signals at a frequency higher than the critical sampling frequency is compensated by the efficiency of the hardware architecture, which has a filterbank coprocessor dedicated to performing subband decomposition of the input signals.

In the following sections, we first present a description of this DSP architecture. We then describe the adaptive noise canceller structure. Finally, a conclusion of the research and the future work is presented.

2. THE DSP SYSTEM

Figure 1 shows a block diagram of the DSP system [6,7]. The DSP portion consists of three major components: a weighted overlap-add (WOLA) filterbank coprocessor, a 16-bit block-floating point DSP core, and an input-output processor (IOP). The DSP core, WOLA coprocessor, and IOP run in parallel and communicate through shared memory. The parallel operation of

the system allows for the implementation of complex signal processing algorithms in low-resource environments with low system clock rates. The system is especially efficient for subband processing since the configurable WOLA coprocessor splits the fullband input signals into subbands, leaving the core free to do the adaptive processing on the subband signals.

The core has access to two 4-kword data memory spaces, and another 12-kword memory space used for both program and data. The core provides 1 MIPS/MHz operation and has a maximum clock rate of 4 MHz at 1 volt. At 1.8 volts, 30 MHz operation is also possible. The system operates on 1 volt (i.e., from a single battery). With a system clock rate of 1.28 MHz, it consumes less than 1 mW of power.

The system is implemented on two ASICs. A separate off-the-shelf E^2PROM provides the non-volatile storage. The chipset can be packaged into a 6.5 x 3.5 x 2.5 mm hybrid circuit.

The system is clocked at a rate of 5.12 MHz for this application. The sampling rate is 16 kHz. Power consumption is 2.1 mW.

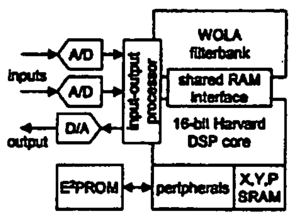


Figure 1: The DSP system block diagram

3. SUBBAND ADAPTIVE NOISE CANCELLATION

The SAF system is implemented on DSP system described in Section 2 The adaptive noise cancellation algorithm uses a 16-band stereo configuration of the WOLA filterbank, with an oversampling factor of 2 or 4. For many applications the low group delay requirement does not allow long analysis time-windows. Consequently, high oversampling factors are used to minimize the aliasing distortion found in systems with critical sampling or low oversampling. This results in near-orthogonal subbands, where energy leakage between adjacent bands is small. As a result, prototype filter design constraints become less stringent. As discussed in [6,7], wide gain adjustment of the subband signals leads to considerable distortion in filterbanks with low oversampling ratios. However, it is quite feasible for the WOLA filterbank to apply wide gain adjustment without generating audible distortions.

Figure 2 shows a block diagram of the subband adaptive noise canceller. The system has two inputs: one for the primary signal (voice from speaker with interfering noise), and one for the reference signal (noise only). The signals are received from microphones that are physically placed for good separation of the signals, but not so far apart as to make the transfer function between microphones too complex to be modeled by the adaptive system. For a headset with a boom, the speech microphone is placed close to the speaker's mouth on the inside of the boom and the reference microphone is placed on the opposite side of the boom facing away from the speaker. Each input signal is passed through the analysis filterbank and split into uniform subbands. The analysis filterbank efficiently decimates the subband signals. The subband processing blocks cancel the noise in the output signal by using a variant of the LMS algorithm that is described in Section 3.2. The subband processing blocks are shown in detail in Figure 3.

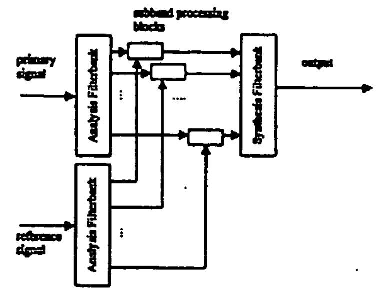


Figure 2: Subband adaptive noise canceller

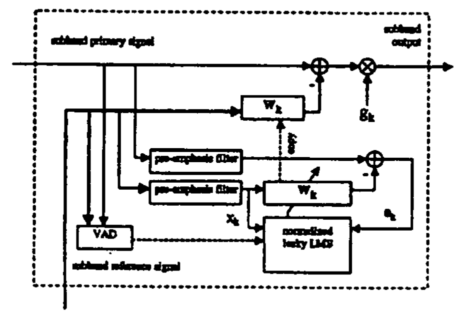


Figure 3: Subband processing block for adaptive noise canceller

3.1. Pre-emphasis Filters

The oversampled input signals received by the subband processing blocks are no longer white in spectrum. In fact, for oversampling factors of 2 and 4, their bandwidth will be limited to $\pi/2$ and $\pi/4$ respectively. As a result, one would expect a slow convergence rate due to eigenvalue spread problem [2]. On the other hand, while the oversampled subband signals are not white, their spectra are colored in a predicable way and can therefore be modified by fixed filters to "whiten" them in order to increase the convergence rate. Thus, the inherent benefit of decreased spectral dynamics resulting from subband decomposition is not lost due to oversampling.

Figure 4 shows a simplified representation of the subband spectra corresponding to white noise input into the filterbank, for a 4-times oversampled configuration. The dashed line shows the spectrum without pre-emphasis. As shown, nearly all the signal power is in the lower quarter of the spectrum. The signal power present in the upper three quarters of the spectrum is

decided by the frequency response of the filterbank's prototype low-pass analysis filter.

We employ a pre-emphasis filter for each subband to amplify the low-level signal components in the high three quarters of the spectrum to flatten the spectrum, thereby reducing the signal's autocorrelation matrix eigenvalue spread, and increasing convergence rate. Figure 5 shows the frequency response of a typical pre-emphasis filter employed in the system. The solid line in Figure 4 corresponds to the spectrum of the subband signal after pre-emphasis. The emphasized subband signals are used only for improving the convergence characteristics of the adaptive filters. As shown in Figure 3, in each subband, the adaptive filter coefficients are copied to a mirror filter that processes the non-emphasized version of the signal to obtain the noise-cancelled signal for synthesis.

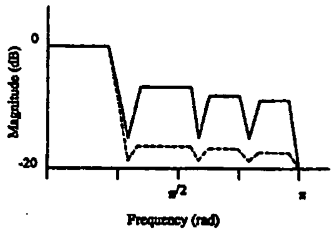


Figure 4: Simplified subband spectrum before pre-emphasis (dashed line) and after pre-emphasis (solid line)

Figure 6 illustrates the change in convergence using a long sequence of white noise input samples into the 16-band WOLA filterbank using an oversampling factor of 4. MATLAB simulations are run with a known finite impulse response system in place to simulate the transfer function between two microphones. The LMS filter mean-squared error (MSE) is the averaged squared difference between the 5 adaptive filter coefficients and the known optimum solution. This value is normalized such that the initial zero values of the adaptive coefficients corresponds to a MSE of 0 dB. The normalized filter MSE is then averaged across the 16 subbands. Note that Figure 6 merely illustrates the difference in average MSE for the finite input sequence; both systems will ultimately converge to the same steady state solution.

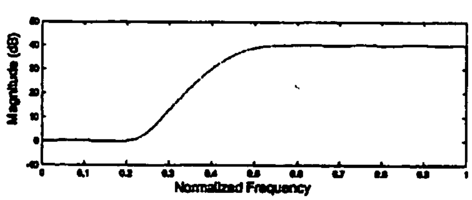


Figure 5: Pre-emphasis filter response

3.2. Subband Adaptation Algorithm

The filter in the kth subband, $\mathbf{w}_k$, is adapted according to equation (1), where n is the time index, μ_k is the LMS step-size parameter, e_k is the error signal, L is the adaptive filter length, $\mathbf{x}_k$ is a vector containing the last L complex samples of emphasized subband reference signal X_k, $\overline{\sigma_k}^2$ is an estimate of the power of X_k, and ε is a small constant used to avoid division by zero. The normalized and "leaky" variant of the complex LMS algorithm is chosen to ensure stability and convergence to a unique solution [8].

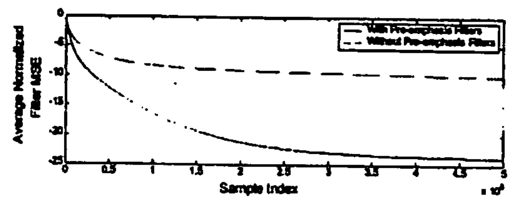


Figure 6: Effect of pre-emphasis filter on adaptive filter convergence

$$\mathbf{w}_k(n+1) = (1-\gamma\mu_k)\cdot\mathbf{w}_k(n) + \frac{\mu_k \cdot \mathbf{x}_k(n)\cdot e_k{}^*(n)}{L\cdot\overline{\sigma_k}^2(n)+\varepsilon} \quad (1)$$

It is possible to vary the subband LMS parameters such as filter length and LMS step-size parameter μ, independent of parameters of adjacent bands since the bands are almost orthogonal. As described below, we have implemented a system with varying values of μ_k, constant leakage factor γ across all bands, and 5 complex coefficients for each adaptive filter.

The values for μ_k are chosen such that peak noise cancellation in slowly varying noise is achieved within approximately 5 seconds. Faster convergence is possible by increasing μ_k, but it comes at the cost of increased artifacts in the enhanced speech. In bands beyond 4 kHz, the filters are more aggressively adapted using increasing values for μ_k since the higher bands contain less speech energy and therefore there is less distortion introduced by quickly adapting filters.

The leakage factor γ effectively adds white noise to the input signal and ensures convergence to a unique solution [8]. It also allows the filters to re-initialize themselves by slowly leaking to zero in the absence of input X_k. γ is chosen such that the factor $(1 - \gamma\,\mu_k)$ is very close to 1. This keeps the filter coefficient bias created by using leaky LMS to an acceptable value, while still adding some whitening effect.

The filter length is chosen as a compromise between computational requirements and the system's ability to model the physical system between primary and reference microphones. Filters that are too long will use up all available processing power and will lead to slow convergence. Filters that are too short will result in a truncated model of the system between microphones, and therefore limit the degree of noise cancellation. Since the adaptive filters in our system operate in a decimated domain and are comprised of complex coefficients, they combine to model a fullband system with a comparably more complex response. The 5 complex coefficients per adaptive filter provide adequate modeling capability, while conserving processing resources.

The existence of multiple filters allows the filter updating to be interleaved across successive time slots for efficiency. For

10-4

example, grouping the subbands into 2 groups of 8, then updating alternate groups at every time slot reduces the computational requirements per time slot by a factor of 2. The power estimate $\bar{\sigma}_k^2$ is calculated using a first-order IIR smoothing filter with a time constant of approximately 1 ms.

The constant gain factors g_k (see Figure 3) are used to scale the noise-cancelled signal before it reaches the subband output and subsequently enters the synthesis stage. We have found that the undesirable leakage of the speech signal into the reference signal in practical systems causes some inadvertent cancellation of speech, particularly in the low frequencies. The static gain factors are set to compensate for this mild low frequency loss. Also, in real-time hardware implementation (reported in Section 4), these gains can be used for microphone equalization.

An optional voice activity detector (VAD) freezes the adaptation of the filters when speech is present. The VAD is particularly useful in physical configurations where microphones are placed such that the speech signal easily leaks into the reference signal. The contamination of the reference signal hinders convergence of the filters. This is avoided by allowing the filters to adapt only when the VAD has detected a pause in speech. The VAD calculates the power in a low band-group and a high band-group. It tracks the changes in the ratio of these powers in order to detect the presence of speech in the primary signal. It is designed to have a bias towards over-detection (false alarms) rather than under-detection (missed speech). A hangover counter is used to prevent the misclassification of trailing portions of speech as noise or silence, thereby improving the reliability of pause detection. Testing shows that activation of the VAD slows down the convergence but does not affect the degree of noise cancellation achieved after convergence.

4. PERFORMANCE EVALUATION

Off-line evaluation tests have been completed for various types of noise (white, pink, car, airplane, babble, and similar noises) in the presence of speech. Table 1 shows the results of a comparison of simulated fullband (128-coefficient FIR) and subband (16 x 8-coeffecient FIR) systems using the same input length. The primary input has a 0 dB signal-to-noise ratio (SNR) with no speech leakage to the reference input. The algorithm parameters (filter length, μ_k and γ) are chosen for each system such that SNR improvement in white noise is similar. The results illustrate how the subband implementation performs consistently for various noise conditions, while the fullband implementation does not. As evident from the table, the proposed SAF has a superior performance for both non-stationary (like babble noise) and colored noises (like pink noise) due to the whitening effect of the SAF system and a faster convergence. Informal listening shows very little audible distortion of speech.

A real-time version of the proposed SAF system is implemented on the DSP system described in Section 2. The preliminary results using a variety of double-microphone boom-style headsets show an average improvement (for different types of noise with input SNR in 0-5 dB range) in SNR of 10 dB on a real system. This is promising considering the effects of implementation on a 16-bit block-floating-point system using a real headset that permits leakage of speech into the reference microphone.

Table 1: Comparison of simulation results for fullband and subband systems

	SNR improvement (dB)	
	Fullband system	**Subband system**
White noise	25.5	25.7
Pink noise	18.7	25.3
Airplane noise	17.3	23.0
Babble noise	16.4	25.2
Traffic noise	17.4	25.2
Car noise	20.7	25.6

5. CONCLUSIONS AND FUTURE WORK

An SAF noise cancellation system was developed for a highly oversampled filterbank. The system was implemented on an ultra low-resource platform. To improve the convergence rate, we proposed and implemented pre-emphasis filters to improve the performance of the adaptive subband-LMS algorithm. In real-life environments, the noise cancellation system delivers approximately 10 dB reduction of noise power with little distortion of speech, while requiring modest resources in terms of space and power. It performs well in colored noise and shows faster convergence than a fullband implementation. No other system known to the authors delivers such performance with such small size and low power consumption.

Future work will include a complete evaluation of our real-time system and investigation of optimal design criteria for the pre-emphasis filters, as well as alternate means of subband signal whitening. Also, more research can be done to explore the usage of other adaptation strategies on the DSP system.

6. REFERENCES

[1] B. Widrow et al., "Adaptive noise cancellation: Principles and applications". *Proc. IEEE*, vol. 63, no. 12, Dec. 1975.

[2] Haykin, S., *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 3rd Edition, 1996.

[3] A. Gilloire and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments and Applications to Acoustic Echo Cancellation". *IEEE Trans. Signal Processing*, vol. SP-40, no. 8, pp. 1862-1875, Aug. 1992.

[4] J. J. Shynk, "Frequency-Domain and Multirate Adaptive Filtering". *IEEE Signal Processing Magazine*, pp. 14-37, Jan. 1992.

[5] S. Weiss, "On Adaptive Filtering in Oversampled Subbands", PhD. Thesis, Signal Processing Division, University of Strathclyde, Glasgow, May 1998.

[6] R. Brennan and T. Schneider, "Filterbank Structure and Method for Filtering and Separating an Information Signal into Different Bands, Particularly for Audio Signal in Hearing Aids". United States Patent 6,236,731. WO 98/47313. April 16, 1997

10-5

[7] R. Brennan and T. Schneider, "A Flexible Filterbank Structure for Extensive Signal Manipulations in Digital Hearing Aids", *Proc. IEEE Int. Symp. Circuits and Systems*, pp.569-572, 1998.
[8] Hayes, M., *Statistical Digital Signal Processing and Modelling*, John Wiley & Sons, Inc., New York, 1996.

Additional Description C

Technical Report

"Subband Adaptive Signal Processing in an Oversampled Filterbanak"

11-1

Sub-band Adaptiv Signal Processing in an Ov rsampled Filterbank

This technology is applicable for digital signal processing applications where it is desirable to implement an adaptive signal processing algorithm in an oversampled WOLA filterbank.

Subband adaptive signal processing in oversampled filterbanks is applicable in a wide range of technology areas including

- Adaptive noise reduction algorithms
- Adaptive directional signal processing with microphone arrays
- Feedback reduction for hearing aids
- Acoustic echo cancellation

A common approach in the signal processing applications listed above is to use a time domain approach, where a filterbank is not used, and a single adaptive filter acts on the entire frequency band of interest. This single time domain filter is typically required to be very long, especially when applied to acoustic echo cancellation. Computational requirements are a concern because longer filters require increasingly more processing power (i.e., doubling the filter length increases the processing requirements by *more* than two). Through the use of the oversampled WOLA filterbank, the single time domain filter can be replaced by a plurality of shorter filters, each acting in its own frequency sub-band. The oversampled WOLA filterbank and sub-band filters provide equal or greater signal processing capability compared to the time domain filter they replace – at a fraction of the processing power.

A longer filter typically requires more iterations by its adaptive controlling algorithm to converge to its desired state [Haykin, Simon. *Adaptive Filter Theory*. Prentice Hall. 1996]. In the case of an adaptive noise cancellation algorithm, slow convergence hampers the ability of the system to quickly reduce noise upon activation and to track changes in the noise environment. Thus, utilising the oversampled WOLA filterbank results in faster convergence and improved overall effectiveness of the signal processing application.

Yet another benefit of sub-band adaptive signal processing in an oversampled filterbank is referred to as the "whitening" effect. A white signal has a flat spectrum; a coloured signal has a spectrum that significantly varies with frequency. The WOLA filterbank decomposes coloured input signals into sub-band signals with spectra that are "whiter" than the wide-band signal. Due to oversampling, the whitening effect occurs in only part of the spectrum; however, this behaviour is predictable and uniform across all bands and can therefore be compensated for by emphasis filters (described later). The commonly used least-mean-square (LMS) algorithm for adaptive signal processing performs best with white signals. Thus, the whitening effect provides a more ideally conditioned signal, improving system performance.

Yet another benefit of sub-band adaptive signal processing in an oversampled filterbank is th ability to set varying algorithm parameters for individual frequency bands. For example, a noise cancellation algorithm can have filters that are set up to converge at different rates for different sub-bands. In addition, the adaptive filters can have different lengths. The increased number of possible parameters allows the system to be more effectively tuned according to the requirements of the application.

In situations in which processing power is limited or must be conserved, the update of the adaptive filter groups can be interleaved. Thus, an adaptive filter is occasionally skipped in the update process but still gets updates at periodic intervals. The processing time required to update a single time domain filter cannot be split across time periods in this way.

In summary, the problems with time domain adaptive signal processing are:

- Long filters required – cannot interleave the update of multiple filters
- Slower filter convergence due to longer filter length
- Performance problems in coloured noise
- Inability to set varying algorithm parameters for individual frequency bands

The oversampled WOLA filterbank also address the problems with traditional FFT-based sub-band adaptive filtering schemes. WOLA filterbank processing was patented for hearing aid applications in US 6,236,731. These problems are:

- Highly overlapped bands that provide poor isolation
- Longer group delay

In addition, oversampled WOLA filterbank processing also provides the following advantages for sub-band adaptive signal processing:

- Programmable power versus group delay trade-off; adjustable oversampling
- Stereo analysis in a single WOLA
- Much greater range of gain adjustment in the bands
- The use of complex gains

An oversampled WOLA filterbank subband adaptive system can also be implemented on ultra low-power, miniature hardware using the system patented by Schneider and Brennan in US 6,240,192.

Some solutions have utilised slight amounts of oversampling possible [M. Sandrock, S. Shmitt. "Realization of an Adaptive Algorithm with Subband Filtering Approach for Acoustic Echo Cancellation in Telecommunication Applications". Proceedings of ICSPAT 2000], but they do not provide the low group delay, flexibility in power versus group delay trade-off and excellent band isolation of oversampled WOLA based adaptive signal processing.

Solutions to problems in time domain adaptive signal processing arising from coloured nois and a long filter are limited. A long filter is often a requirement that is dictated by th particular application, and shortening it would degrade performance. In cases when it is allowable, white noise can be inserted into the signal path to allow the filter to adapt quicker.

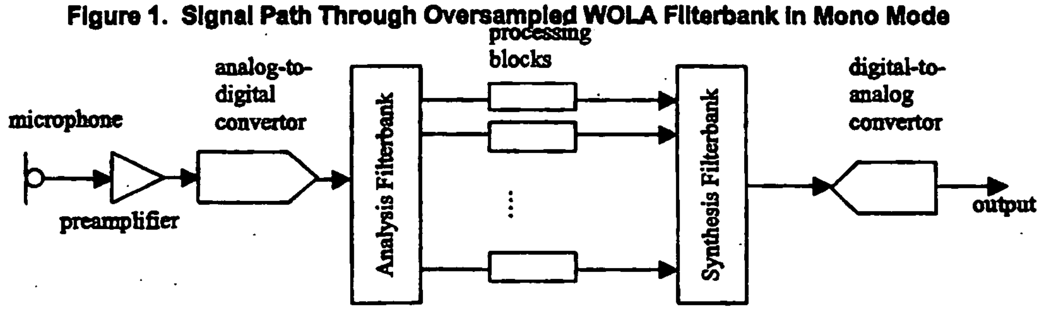
Slow convergence is usually dealt with by choosing algorithm parameters that result in fast convergence while still guaranteeing filter stability. In the LMS algorithm, this is done by increasing the step-size parameter (mu). However, this approach causes considerable distortion in the processed outout signal due to the larger fluctuations of the adaptive filter resulting from a high mu value.

A method used to increase computational speed in time domain signal processing is to perform operations in the Fourier transform domain [J. J. Shynk, "Frequency Domain and Multirate Adaptive Filtering", IEEE Signal Processing Magazine, vol. 9, no. 1, pp.15-37, Jan 1992]. A section of the signal is transformed, operated on, then undergoes an inverse transformation. Methods are well known for performing specific operations in the transform domain that directly correspond to linear convolution (a common operation) in the time domain, but require less processing time. The added requirement of having to calculate the Fourier transform and inverse Fourier transform is offset when the signal can be transformed in blocks that are sufficiently large.

Adaptive signal processing in an oversampled WOLA filterbank provides

- very low group delay
- a flexible power versus group delay tradeoff
- highly isolated frequency bands
- wide-ranging band gain adjustments
- variable algorithm parameters in different sub-bands: filter length, convergence rate, etc; algorithm parameters can be optimally adjusted to meet computation as well as other performance constraints
- faster convergence of adaptive filters
- reduced computation time
- improved performance in coloured noise
- ability to split computational load associated with updating adaptive filters across multiple time periods

Figure 1. Signal Path Through Oversampled WOLA Filterbank in Mono Mode



11-4

Figure 1 shows the signal path through the oversampled WOLA filterbank operating in mono mode. Figure 2 shows the signal path through th oversampled WOLA filterbank operating in stereo mode. The logic contained in the processing blocks is dependent on the particular application. For sub-band adaptive signal processing, these blocks contain adaptive filters and their associated control logic.

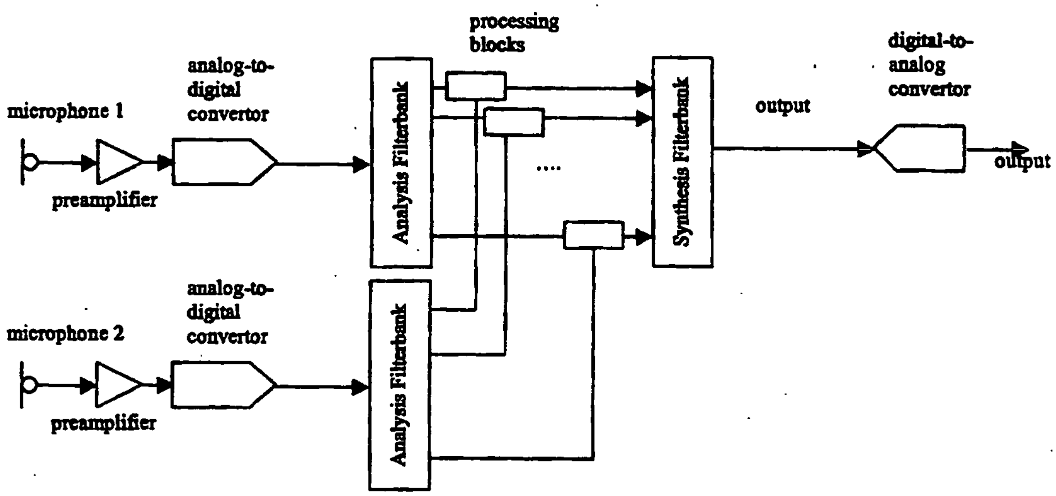


Figure 2. Signal Path Through Oversampled WOLA Filterbank in Stereo Mode

The type of filters (recursive or non-recursive), method of controlling the adaptive filters, and number of inputs (one or many) can vary. The LMS algorithm and its variants are widely used in adaptive signal processing for their relative simplicity and effectiveness. Many applications use the two-input stereo configuration, but sub-band adaptive signal processing with one or many inputs is also within the scope of this invention. Furthermore, this invention is not bound to any particular configuration of the oversampled WOLA filterbank (i.e., number of sub-bands, sampling rate, window length, etc).

The WOLA filterbank provides an input to each sub-band block that is highly isolated in frequency. The sub-bands may have independent adaptive parameters, or they may be grouped into larger frequency bands and share properties.

After adaptive processing, the sub-band signals are sent to the synthesis filterbank, wh re th y are recombined to a single output signal. The net effect of th sub-band adaptive filters on this output signal is equal to a singl tim domain filter that is much longer than any one of the sub-band filters.

See US 6,236,731 for a thorough description of WOLA filterbank signal processing.

A description of two main embodiments of this invention follows. Both embodiments are described for noise cancellation applications. This is a typical application of adaptive oversampled WOLA processing, but there are many others. First is a sub-band noise cancellation algorithm that uses a variant of the LMS algorithm, and the oversampled WOLA filterbank in stereo mode. Then another embodiment will be described that also performs noise reduction with a two-microphone configuration and an alternative method for deriving the adaptive coefficients.

Two-microphone LMS Noise Cancellation

Although least-mean squares signal processing is described here, other techniques well known in the art are also applicable. For example, recursive least squares could also be used.

Description

This is an algorithm that is used to cancel the noise in transmitted speech when the speaker is in a noise environment. The listener, not the speaker, experiences the improvement in signal quality. Examples of where is algorithm can be used is telephone handsets, and boom-microphone headsets.

The basic structures used in this algorithm can be applied to other applications as well. One skilled in the art could modify this algorithm for acoustic echo cancellation or acoustic feedback cancellation.

This algorithm is useful for all headset styles that use two microphones for speech transmission.

How It Works

Two-microphone adaptive noise cancellation works on the premise that one signal contains noise alone, and the other signal contains the desired signal (speech) plus noise that is correlated with the noise in the first signal. The adaptive processing acts to remove the correlated elements of the two signals. Since the noise signals are (assumed to be) correlated and the speech is not, the noise is removed.

11-6

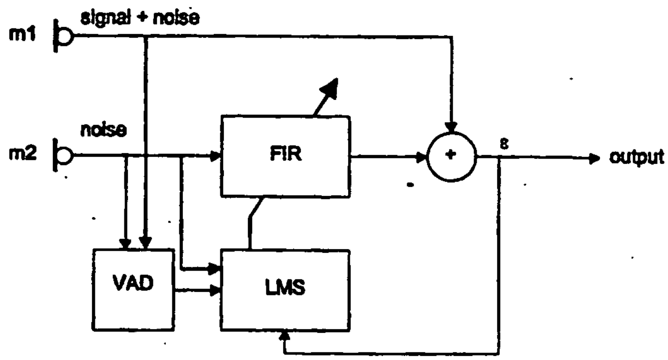


Figure 3. Time-domain adaptive noise cancellation

Figure 3 shows a block diagram of a time-domain, two microphone adaptive noise cancellation. The LMS block controls the adaptive finite impulse response (FIR) filter in order to minimize the noise appearing at the output. A voice activity detector (VAD) is used to stop or slow adaptation when speech is present. This reduces artifacts in the output signal that are caused by misadjustments of the FIR filter due to the presence of speech. The VAD can use both signals as inputs and employ the differential level as an indicator that speech is present (it is assumed that the m1, the mic facing the talker, will receive a higher level signal that m2). In a headset application, the two microphones could be located on a boom with *m1* facing in and *m2* facing out.

Note that this algorithm can also be implemented in the frequency domain (Figure 4). In this version of the algorithm the processing is done in N bands, each with a complex output signal (magnitude and phase). Again, a VAD is used to stop or slow the adaptation when speech is present. In theory, a frequency domain implementation will offer better performance than a time-domain implementation because it will converge faster and effectively implement longer adaptive filters (which can use interleaved or decimated updates to reduce the computational load). Also, noise rejection for frequency-localized noise is likely to be better.

11-7

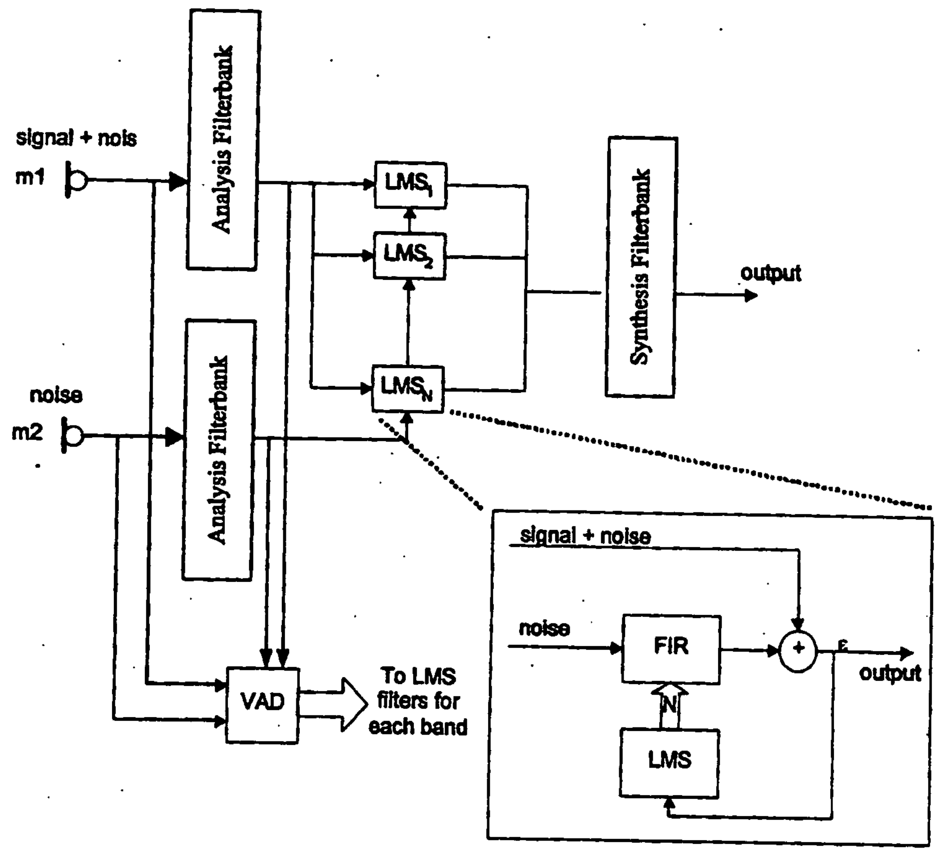


Figure 4. Frequency-domain adaptive noise cancellation

The LMS blocks implement what is well known in the art as leaky normalized LMS. The LMS step-size varies in each sub-band; lower sub-bands contain high speech content and have a smaller step-size, while higher sub-bands can be more aggressively adapted with a larger step-size due to relatively low speech content.

A key addition to the leaky normalized LMS algorithm is the use of a spectral emphasis filter. This additional filter is static and serves to whiten the LMS input signals for faster convergence. Oversampling in the filterbank inherently produces sub-band signals that are coloured in a predictable way. In the case of two times oversampling, the bottom half of the sub-band spectrum has relatively high energy and is relatively flat compared to the upper half of the spectrum, which contains very little energy. The spectral emphasis filters amplify the part of the spectrum known to have lower energy, thus the signal is modified towards the ideal case of being white.

11-8

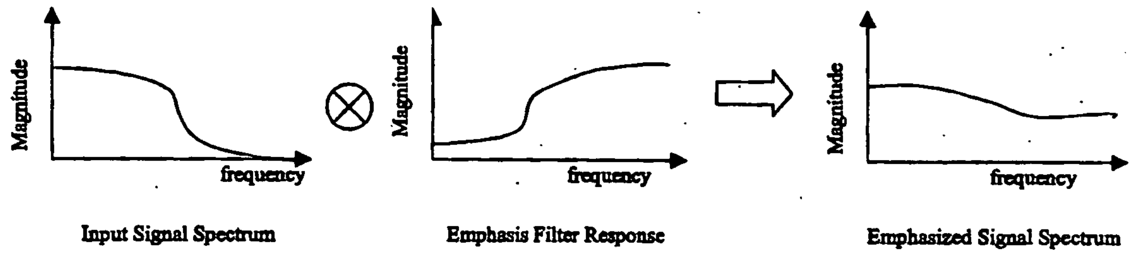


Figure 5. Illustration of Spectral Emphasis

Figure 5 illustrates the spectral emphasis operation. The oversampled input signal has a drop off in energy towards high frequencies, and the emphasis filter is designed to amplify the high frequencies. The filtering operation results in a signal spectrum that is flatter, or whitened.

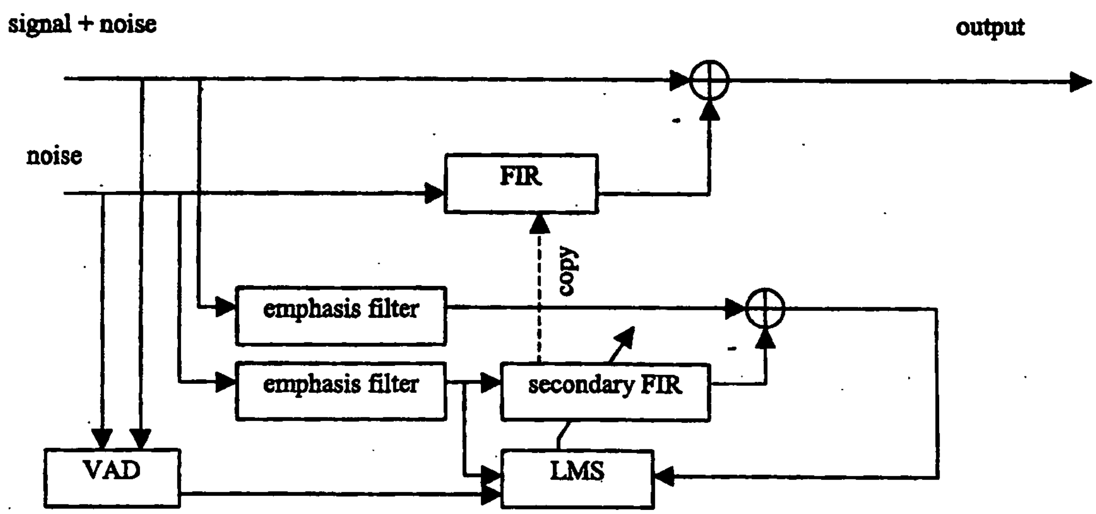


Figure 6. LMS_N Block with Spectral Emphasis Filter

Figure 6 shows the signal flow of the LMS_N block when the spectral emphasis filter is used. Both the signal plus noise and the noise only inputs are filtered and whitened before they are used by the LMS block to update the secondary FIR filter. Since the output signal generated using the secondary filter has been affected by the emphasis filter, it is not suitable to be sent to the synthesis filterbank. It is not desirable to have a synthesis filterbank output signal that has been noticably emphasized in some frequency regions. To void this, a copy of the secondary FIR is used to operate on the unemphasized signals to generate the signal to be synthesized.

The design of the mphasis filter is dependent on the oversampling factor us d in the WOLA filterbank. Given the oversampled WOLA filterbank parameters,

the spectral properties of the sub-band signals can be determined, and an appropriate emphasis filter can be designed. It can be implemented as a FIR filter or an IIR (infinite impulse response) filter.

Two-Microphone Wiener Noise Reduction

Description

This is a transmit algorithm that uses a block-based interference cancellation scheme similar to the *Two-Microphone LMS Noise Cancellation* algorithm. The basic technique is Wiener noise reduction [B. Widrow, S. Stearns. *Adaptive Signal Processing*. Prentice Hall. 1985]. A completed and "tuned" version of this algorithm is likely to provide performance similar to the Clarity algorithm (http://www.claritycom.com/). This algorithm is new for Dspfactory and should be considered a research project since we have no experience with two-microphone Wiener algorithms (however, we do have significant experience with signal-microphone Wiener noise reduction).

This algorithm is useful for all headset styles that use two microphones for speech transmission.

How It Works

This algorithm utilizes the stereo processing mode of the WOLA filterbank. Two signal are simultaneously transformed to the frequency domain: one signal is *speech + noise*, the other is *noise* alone. The processing acts to remove the noise that is correlated between the two signals. Figure 7 shows a block diagram of this processing.

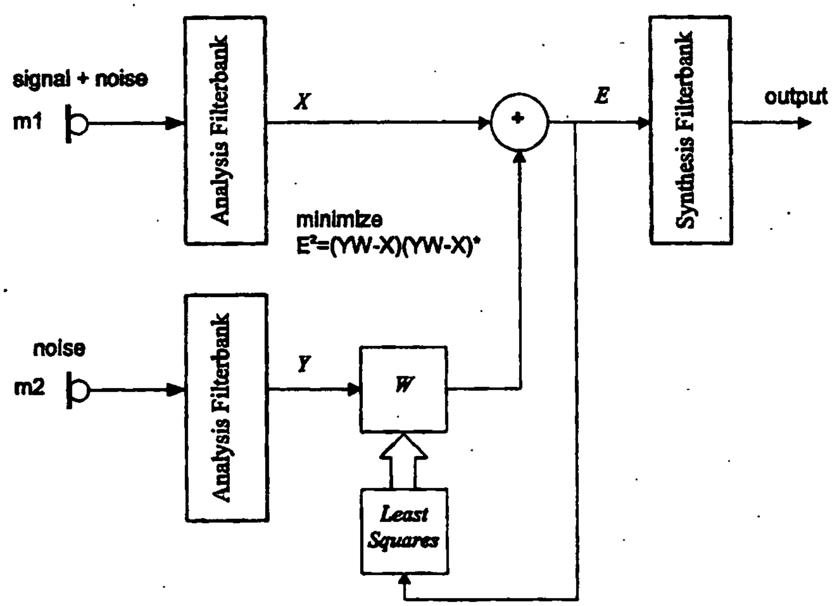


Figure 7. Two-microphone Wiener noise reduction

The solution that minimizes E^2 is the equation:

11-10

$$W=r_{xy}/R_{x} \qquad (1)$$

where R_x is the auto-correlation matrix of X and r_{xy} is the cross-correlation matrix of X and Y [M. H. Hay s. *Statistical Digital Signal Processing and Mod ling*. John Wiley & Sons, Inc. 1996].

If R_x and r_{xy} are estimated using only the most recent sample of X and Y, the value of adaptive weight W_k at time index n is

$$W_k(n) = Y_k(n) / X_k(n),$$

where k is the sub-band index.

Thus, update of an adaptive weight only requires division of the complex values $Y_k(n)$ and $X_k(n)$. Taking one-sample estimates of the auto-correlation and cross-correlation matrices eliminates the need to perform the matrix inversion of R_x in equation (1).

A novel addition to this algorithm is that of frequency constraints. If left unconstrained, adjacent bands may have very different gains. While this will result in the lowest noise level (since E^2 will be minimized), it may also result in some undesirable processing artifacts. Constraining the adjustment of the gain vector (*W*), should result in less noise reduction, but fewer artifacts. Equation (2) shows a scheme where the gain in a given band is constrained by the two adjacent bands. Note that this case uses only a single (complex) weight per band. It should be possible to extend this scheme to allow for multiple weights per band. Note that for the single gain case, the matrix is block-diagonal; thus, there are efficient solution methods.

$$\begin{bmatrix} Y_1 & Y_2 & 0 & \dots \\ Y_1 & Y_2 & Y_3 & 0 \\ 0 & Y_2 & Y_3 & Y_4 \\ 0 & 0 & \dots & \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \\ W_k \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \\ X_k \end{bmatrix} \qquad (2)$$

Multi-microphone Wiener algorithms like this have been successfully used for noise reduction in other applications; for example, see *Multi-Channel Spectral Enhancement In a Car Environment Using Wiener Filtering and Spectral Subtraction*, Meyer and Simmer, Proc. ICASSP-97, Vol. 2, pp. 1167-1170.

For further illustration, sub-band adaptive signal processing using the oversampled WOLA filterbank for echo cancellation will be described.

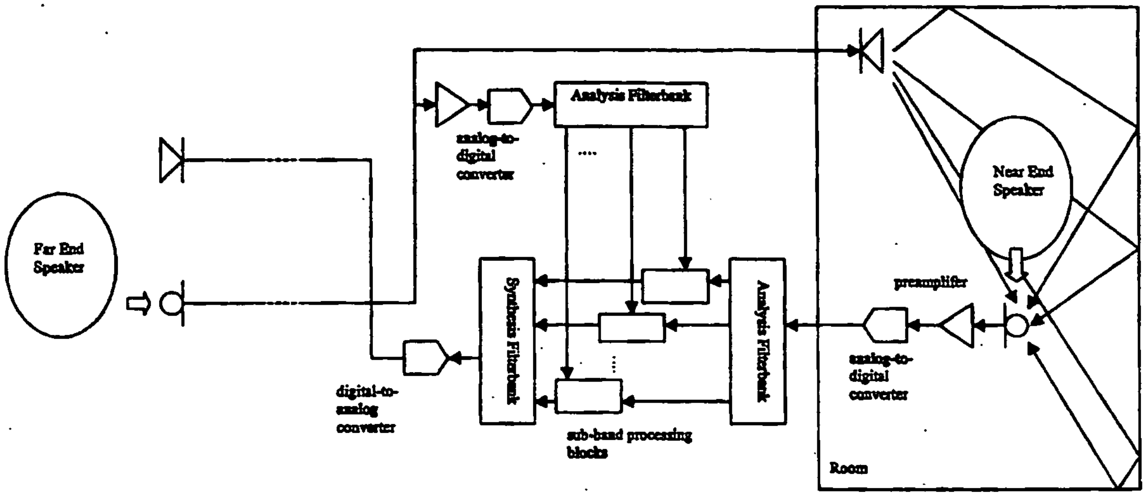


Figure 8. Sub-band Adaptive Acoustic Echo Cancellation with the Oversampled WOLA Filterbank

The goal of acoustic echo cancellation is to remove the far end speaker's voice from the signal that enters the near end microphone and eventually reaches the loudspeaker at the far end (see Figure 8). This allows the near end speaker's voice to be transmitted without echoes of the far end speaker's voice (due to room reverberation), for better intelligibility and less listening effort.

Note that the adaptive signal processing system must deal with a significantly long room response. A single time domain filter will have to contain thousands of coefficients to adequately model this response, and will consequently demand high processing power. Solving this problem using the oversampled WOLA filterbank allows for shorter filters and therefore a savings in processing power over the time domain approach.

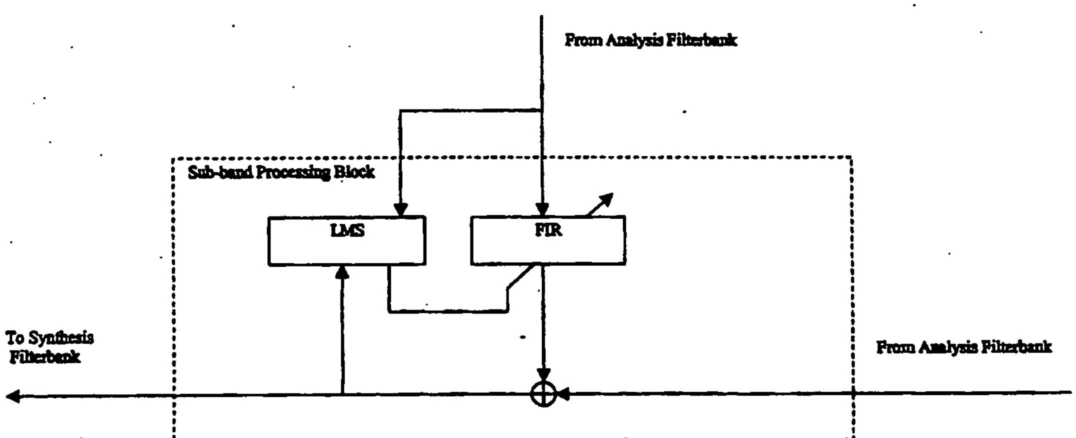


Figure 9. Processing Block for Sub-band Adaptive Acoustic Echo Cancellation with the Oversampled WOLA Filterbank Using LMS

Figure 9 shows the structure of the processing blocks when the LMS algorithm is used to control the adaptive filters. The configuration is much like the noise cancellation system, but the far end speech is considered to be the unwanted noise, and the desired output signal is the near end speech.

The previously described embodiments are examples of adaptive sub-band adaptive signal processing with two inputs. It should be noted that they could be extended to make use of a multiplicity of inputs. A microphone array could be used to capture several input signals, all of which are summed to form the primary (i.e. signal plus noise) signal. Also, in some situations there are several noise sources to be cancelled, therefore a multiplicity of noise censors are required for the reference (i.e. noise) signals.

Details of time domain adaptive algorithms with more than two inputs signals can be found in *Adaptive Signal Processing*, Widrow and Stearns, Prentice-Hall, 1985. The benefits of sub-band adaptive signal processing over time domain adaptive signal processing still hold for these applications. See our co-pending application, "Subband Directional Audio Signal Processing Using an Oversampled Filterbank".

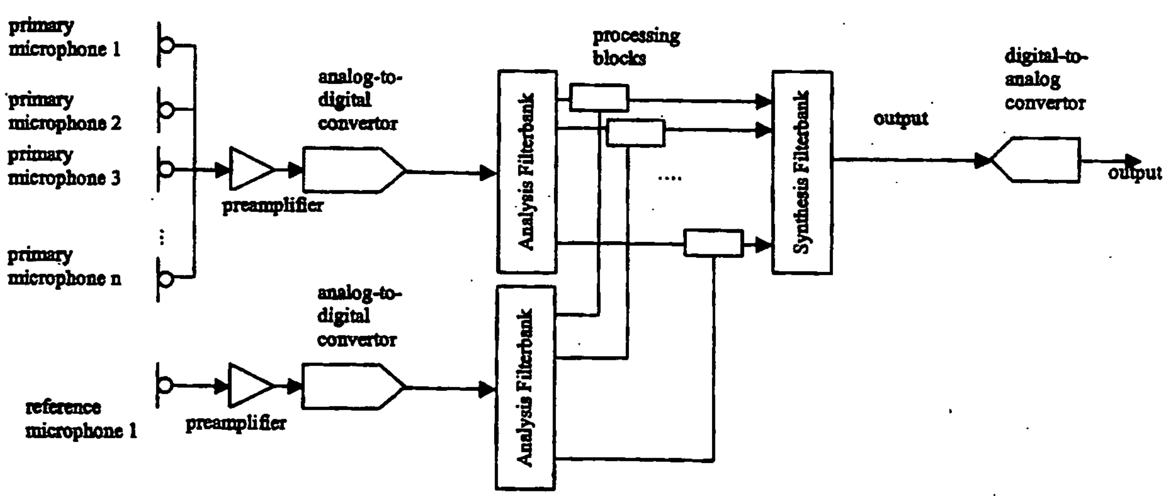


Figure 10. Oversampled WOLA Filterbank Processing Using Microphone Array for Primary Input

Figure 10 illustrates th signal flow for sub-band adaptive algorithm that uses a microphone array for th primary signal.

11-14

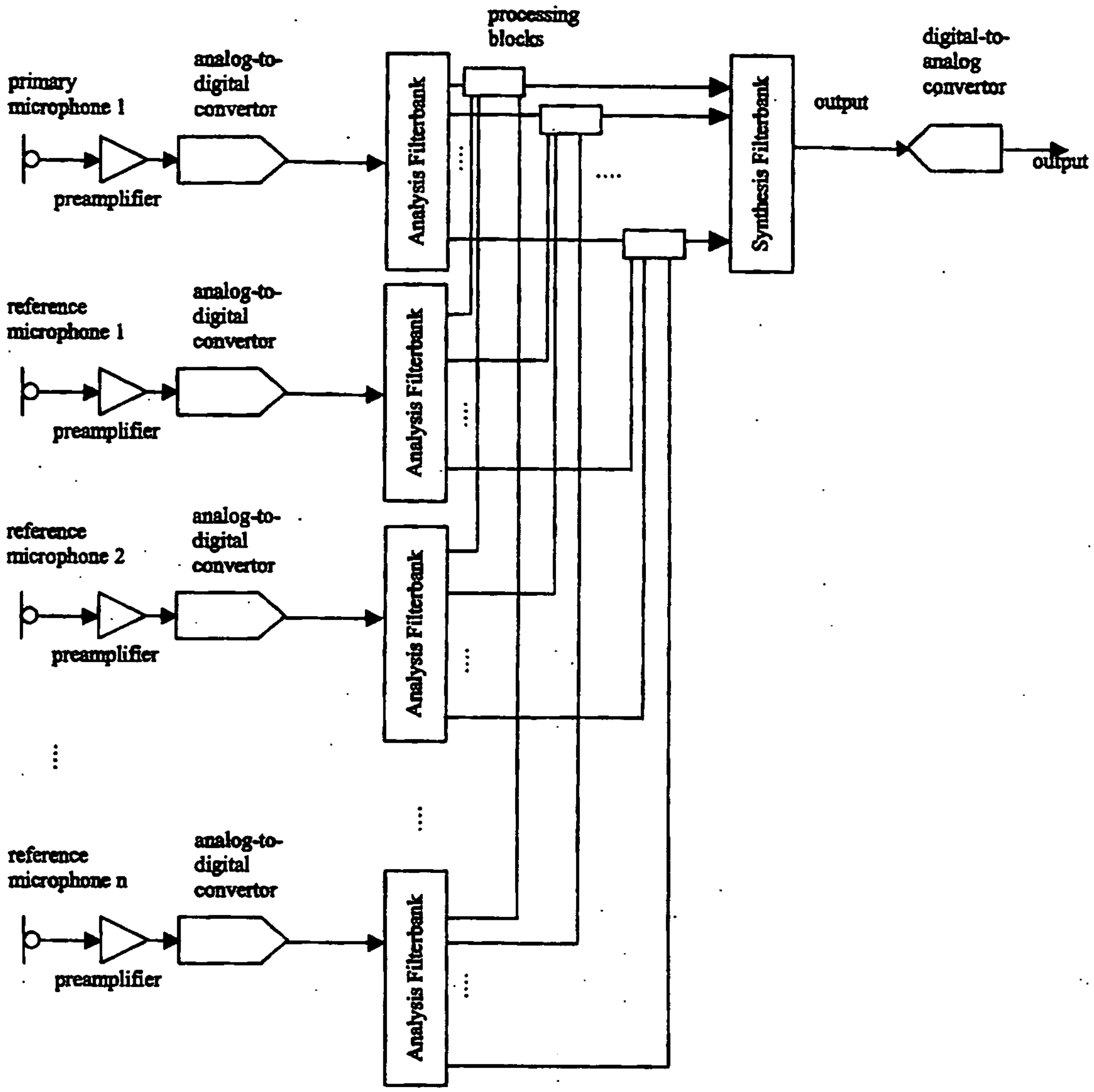


Figure 11. WOLA Filterbank Processing with Multiple Reference Inputs Using LMS

11-15

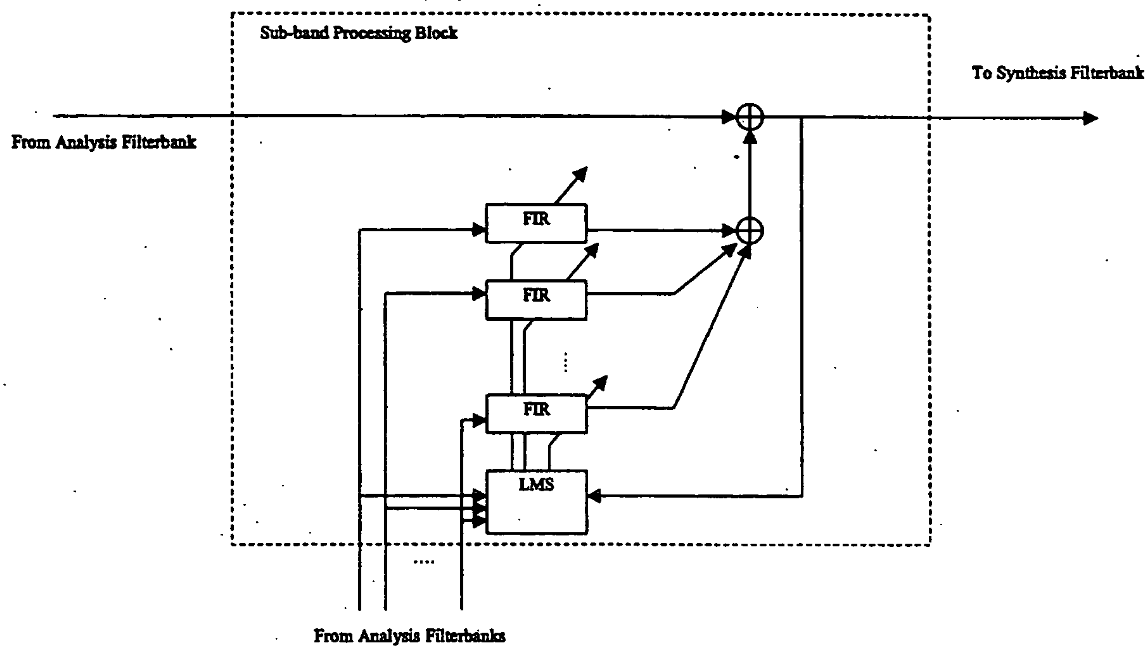


Figure 12. Sub-band Processing Block for WOLA Filterbank Processing with Multiple Reference Inputs Using LMS

Figure 11 illustrates the signal flow for sub-band adaptive algorithm that uses multiple reference microphones and the LMS algorithm. This type of configuration is used in a noise cancellation application when there are more than one noise source. One microphone is used for each noise source to provide a reference signal, which is adaptively filtered and then subtracted from the primary signal (see Figure 12).

What is claimed is:

1. A method of improving the convergence properties of the oversampled subband adaptive filters, the method comprising steps of:

(a) whitening by spectral emphasis, where, after WOLA analysis, subband signals are decimated M/OS where M is the number of filters and OS is the oversampling factor; or

(b) whitening by additive noise, where high-pass noise is added to bandpass signals to make them whiter in spectrum; or

(c) whitening by decimation, where the subband signals are further decimated by a factor of DEC<OS; or

(d) a combination of said steps (a), (b) and (c).

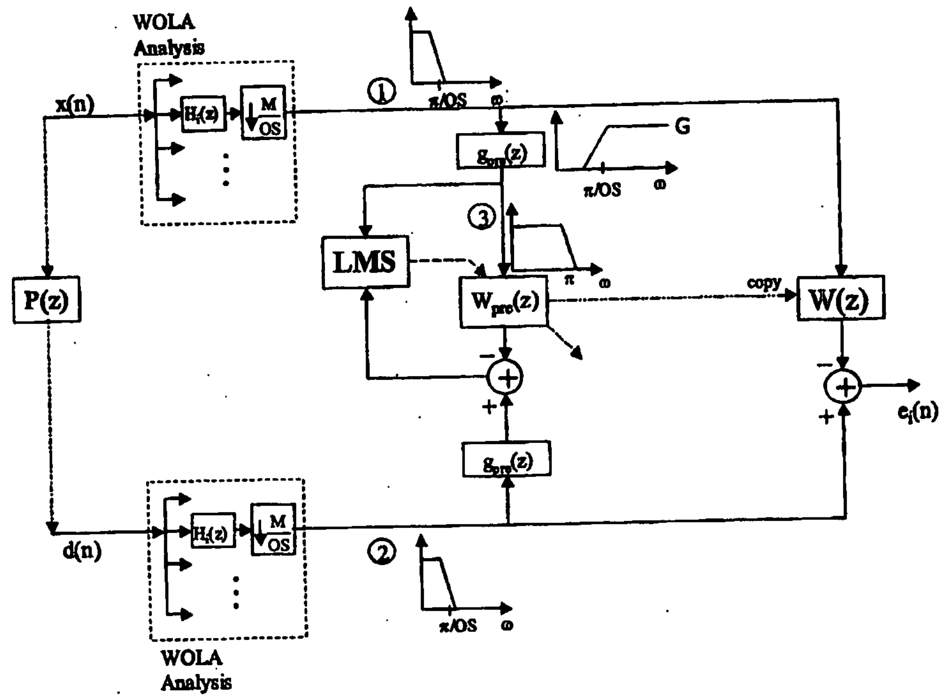


Figure 1: Block diagram of whitening by spectral emphasis method

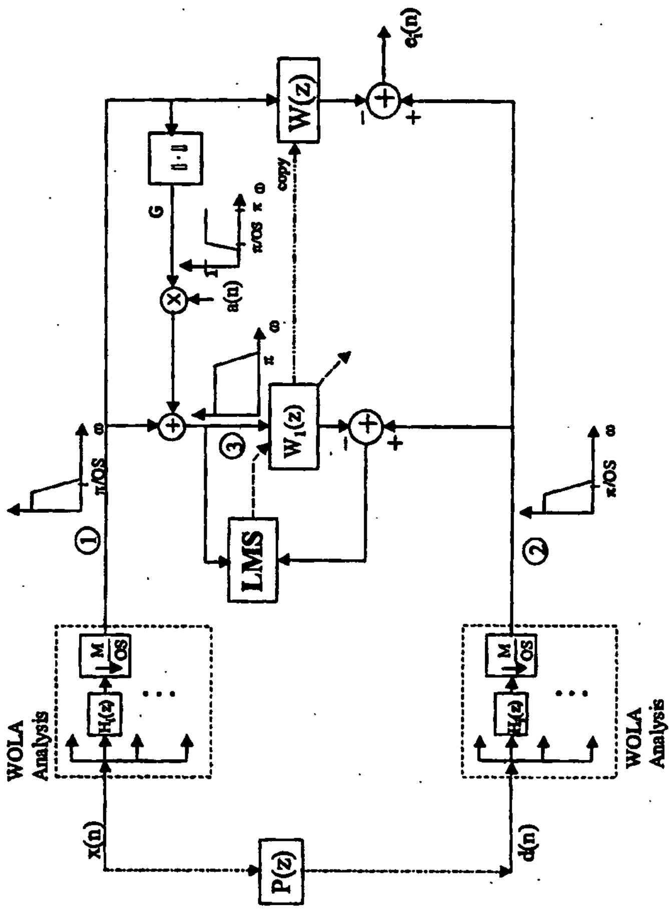


Figure 2: Block diagram of whitening by additive noise method.

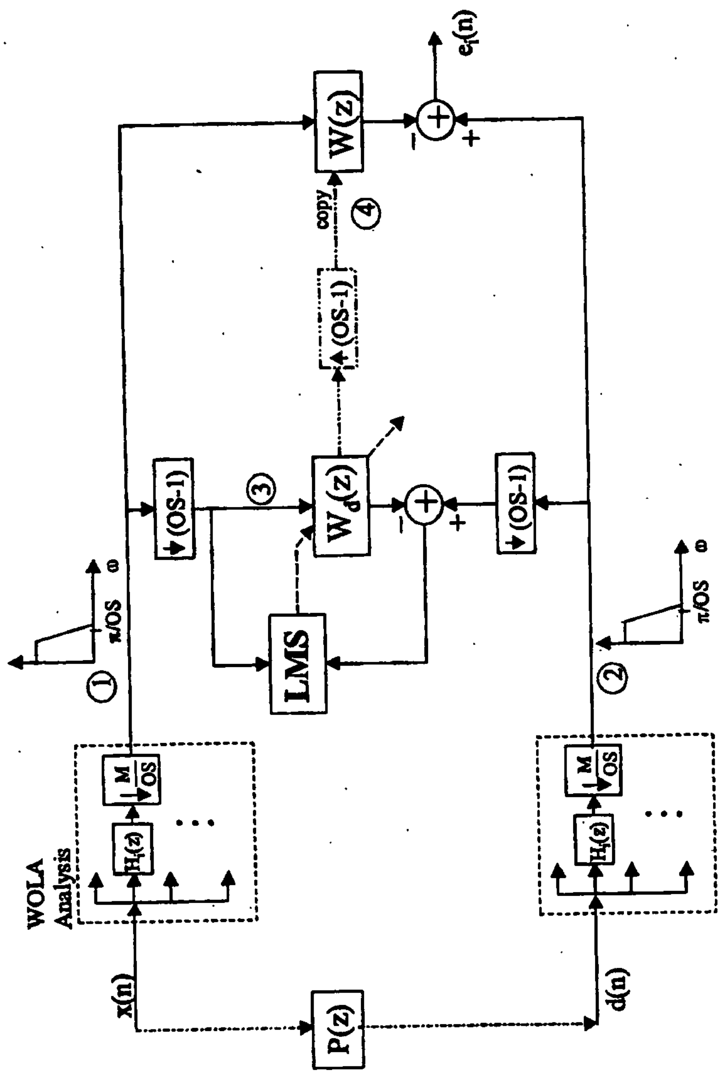


Figure 3: Block diagram of whitening by decimation method.

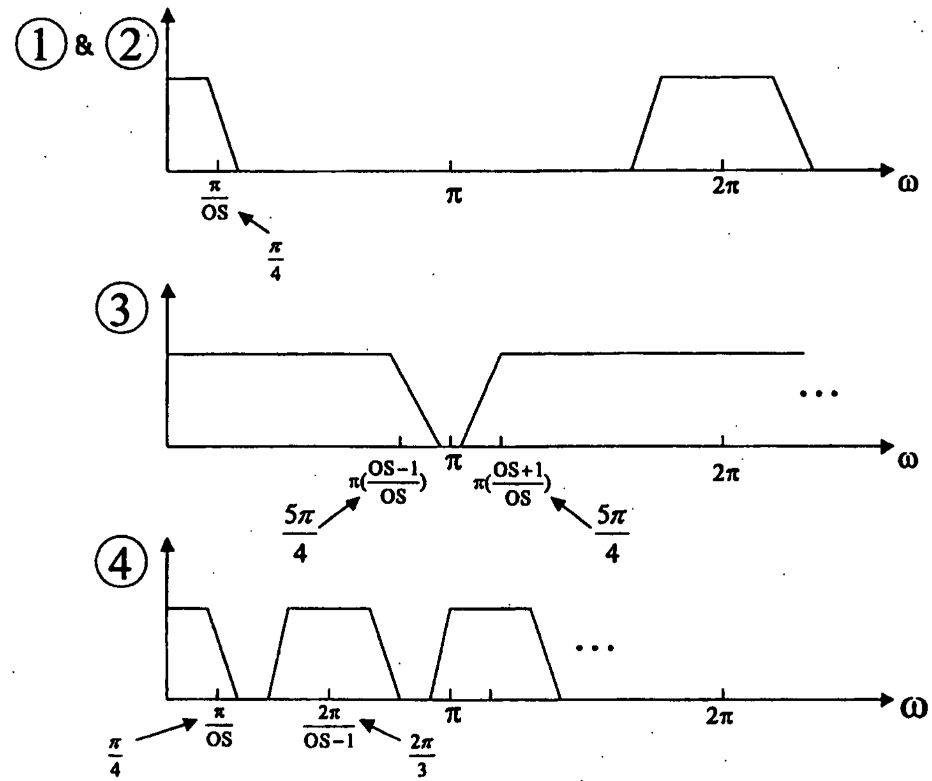


Figure 4: Signal Spectra at various points of Figure 3.

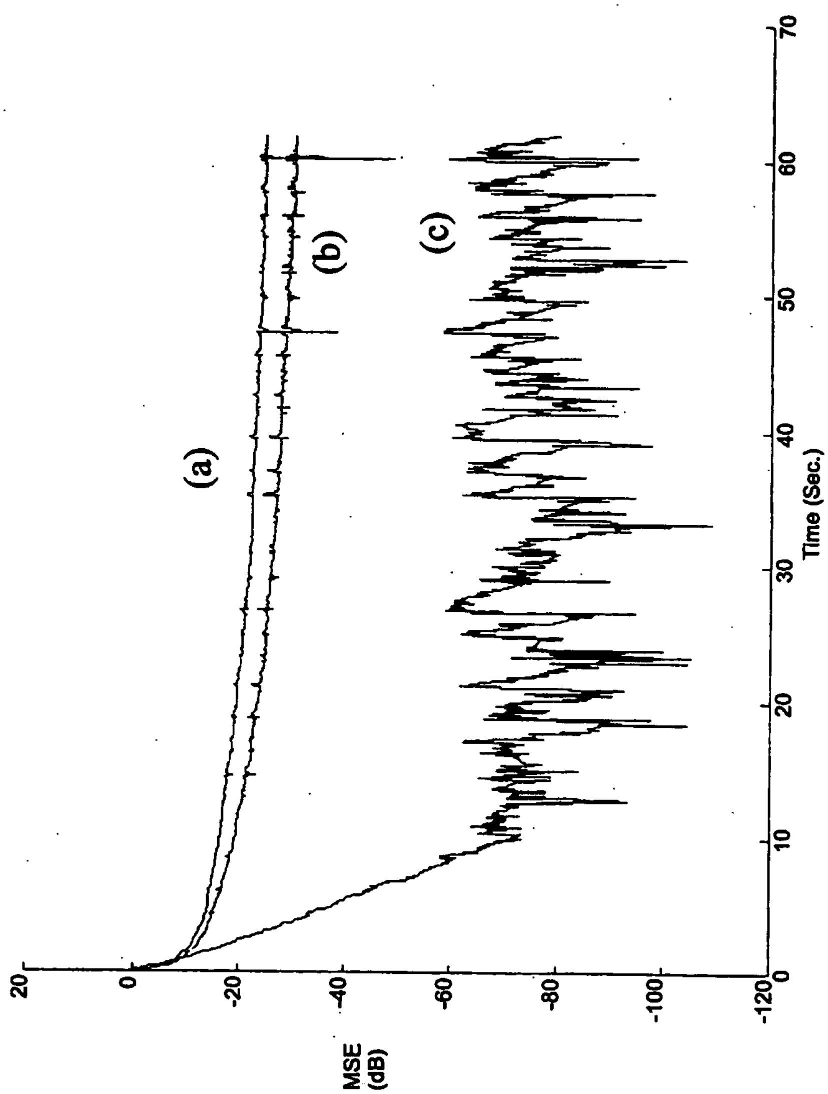


Figure 5: Average Normalized Filter MSE for speech in 0 dB SNR White noise, (a) without whitening, (b) whitening by spectral emphasis, (c) whitening by decimation.

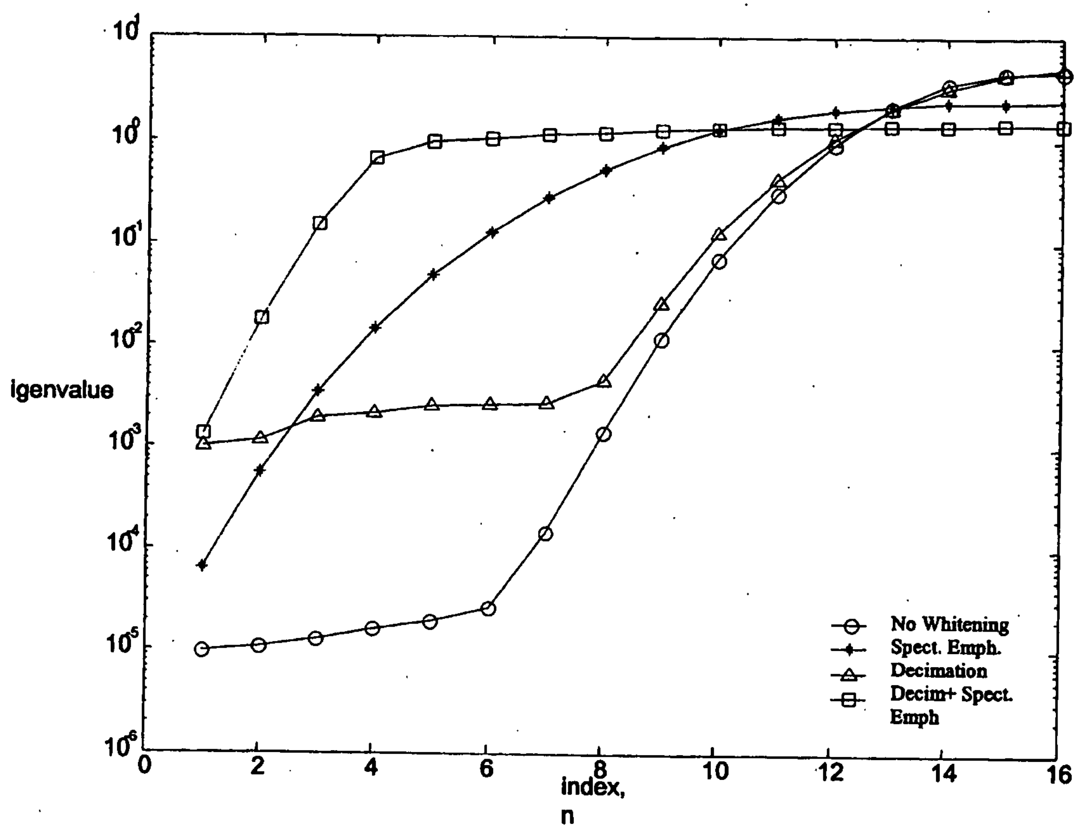


Figure 6: Eigenvalues of the autocorrelation matrix of the reference signal for: No whitening, whitening by spectral emphasis, whitening by decimation, and whitening by decimation and spectral emphasis.

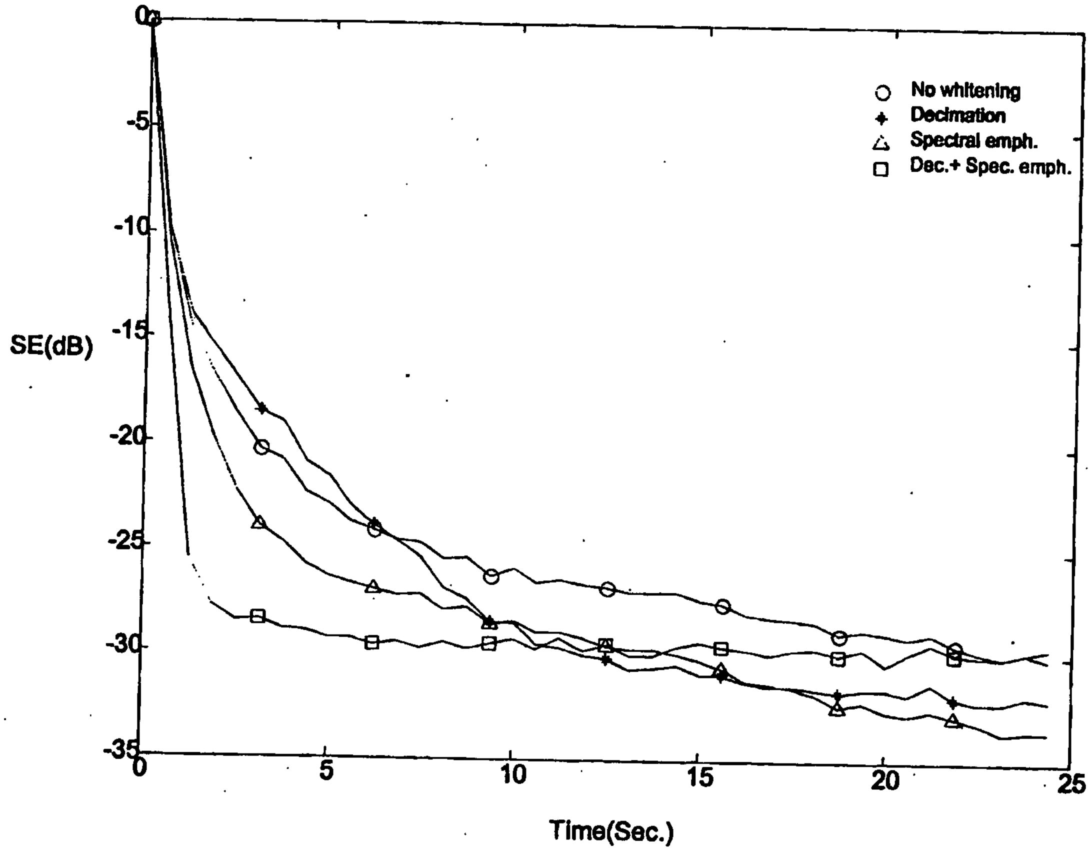


Figure 7: Measured mean-squared error for: No whitening, whitening by spectral emphasis, whitening by decimati n, and whitening by decimation and spectral emphasis.

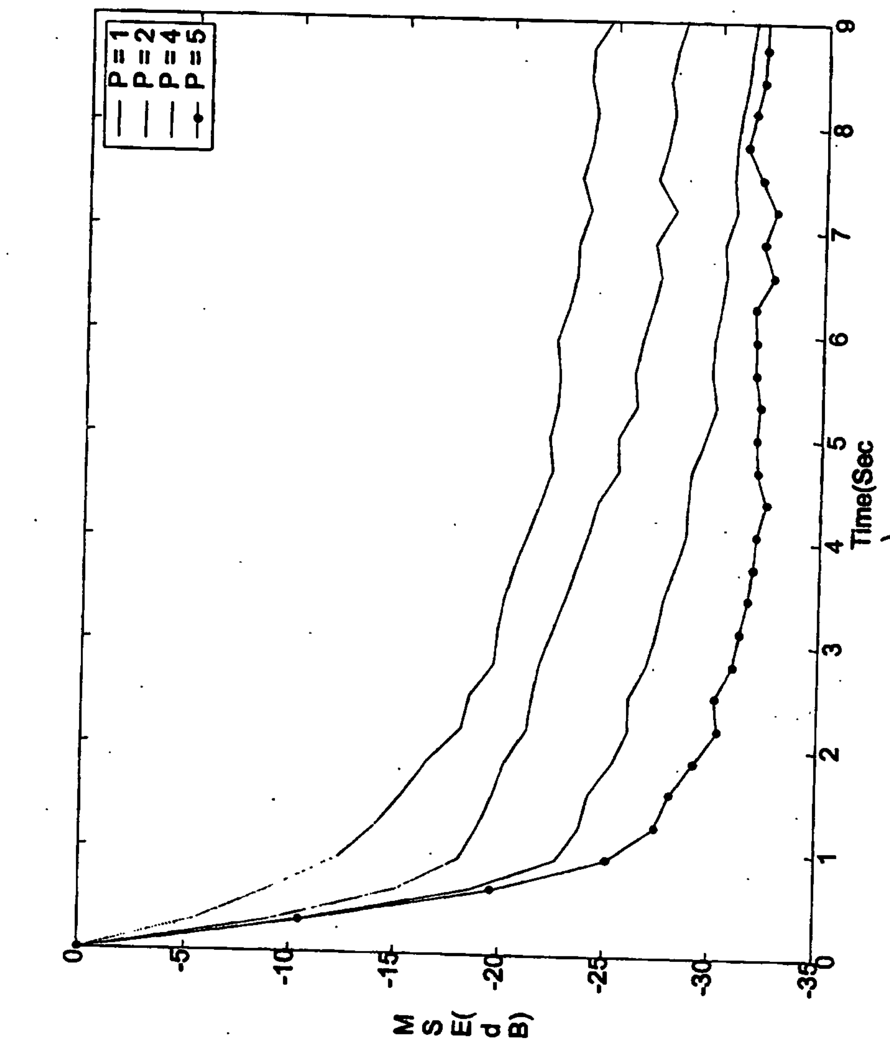


Figure 8: Measured mean-squared error for APA with different orders.